OSU WPL 31 (1985) 64-121

# Syntactic Conditions on Two Types of English Cliticizations in GPSG

Annette Sue Bissantz

## 0. Introduction

Much of the recent work in morphosyntax[1] has focused on characterizing the distinctive properties of clitic elements and the grammatical role of cliticization rules in the languages of the world. Special emphasis has been placed on distinguishing clitic elements from other types of bound morphemes, devising typologies for clitics, and locating rules of cliticization within the grammar as a whole. Though not a necessary feature of such studies, the syntactic framework most often used has been some form of transformational grammar. In this thesis I will look at the phenomenon of cliticization from the point of view of a relatively new theory of syntax, that of Generalized Phrase Structure Grammar[2]. In particular I will examine two forms of English cliticization, Auxiliary Reduction and Complementizer Contraction, which have not received an adequate treatment in transformational grammars and show how they can be accounted for easily and elegantly within the GPSG framework. In a more general vein, I will also show that the nature of syntactic rules in GPSG in part predicts the existence of a separate component within the grammar for cliticization rules; a division independently argued for by many other researchers.

The two types of cliticization I will be interested in here are Auxiliary Reduction (AR) and Complementizer Contraction (CC). AR is responsible for alternations such as the one in (0.1), CC for those like (0.2):

0.1) a. Pita is almost done.
b. Pita's almost done.
0.2) a. Pita wants to get done.
b. Pita wansta get done.

These two rules in particular were chosen for study precisely because they have been the subject of so much discussion in recent linguistic literature. In all the debate surrounding these constructions one can isolate at least two separate issues: 1) what would be the best way of stating the conditions under which AR and CC take place and 2) how should these rules be incorporated into a grammar of English. While some theorists have claimed that the application of AR and CC is dependent upon stress levels in candidate sentences, many others have argued that syntactic structure is the primary determining factor. Furthermore, even within the latter group there has been a great deal of disagreement over how, precisely, to characterize this dependency. Similarly, the rules of AR and CC themselves have been treated differently by different researchers--being sometimes included in the phonology, sometimes in the morphology, and sometimes in the syntax of the language. In the following sections I will attempt to deal with both of these issues.

In section one I will justify my claim that AR and CC are rules of cliticization rather than, for example, simply the result of phonological reductions or affixation processes. I will also show that low stress levels do <u>not</u> guarantee the applicability of these rules and present

preliminary arguments in favor of separating them from other rules in the grammar. In section two I present some of the more well-known formulations of AR and CC and the types of data each can and can not account for. While the main purpose of these discussions is to define the problem at hand, it should be noted that a theory which can deal with these facts in a simple and elegant way would represent a significant improvement over these alternate proposals. Section three contains a brief summary of the basic tenets of GPSG and shows how the correct generalizations about the types of syntactic structures that allow AR and CC fall out automatically in a GPSG treatment of English. This approach is particularly satisfying in that it provides a straightforward account of the dialect variations found with AR and CC. Section three also shows how assuming a GPSG syntax strengthens the conclusions reached in section one concerning the location of cliticization rules in the grammar.

## 1. AR and CC as Cliticizations.

A basic claim of this thesis is that AR and CC are, indeed, rules of cliticization: that is, rules which result in certain free morphemes being realized as bound dependents of other morphemes in a sentence. I wish to make a distinction here between the actual cliticization operation itself and any possible phonological consequences of that operation. As the latter are frequently idiosyncratic, the phonological form of the clitic or the clitic and its host (i.e. the morpheme on which it is dependent) will often have to be specifically listed in the grammar in much the same way as irregular past tense or plural forms. Before I present arguments in favor of this particular view of AR and CC we will need to know a bit more about the nature of clitics, their classifications and associated properties; this is discussed in the following section.

### 1.1 Clitics and Clitic Typologies.

Clitics are a type of bound morpheme found in many languages. They are unusual in that they act in some respects like words and in other respects like affixes, sharing certain properties with each. They are distinct from words in that they cannot usually bear stress and are phonologically dependent on a 'host' element. They can be distinguished from affixes in that they attach to already formed words rather than to roots or stems to make words, they do not necessarily have a close semantic relationship to their host word and, unlike some derivational affixes, they never affect the lexical category of their host[3]. While these are useful criteria for separating clitics as a group from words and affixes, they do not give any insight into the possible subclasses of clitic elements themselves. Many such subgroupings have been proposed. Nida (1946) divides clitics into two groups: those with alternate free forms and those without alternate free forms. Other classifications have focussed more on positioning, with many scholars[4] drawing a distinction between verbal clitics on the one hand and second position (or 2P) clitics on the other. Verbal clitics, as the name implies, attach only to verbs; they also tend to occupy a different position than their free standing counterparts. The object pronoun clitics found in many Romance languages would qualify as verbal clitics. Examples are given below from Spanish and French:

1.1) a. Veo el libro
see-1sg the book
'I see the book'

b. Lo veo
it see-1sg
'I see it'

1.2) a. Je vois Jean
I see-1sg IO Jean
'I see Jean'

b. Je le vois
I him see-1sg
'I see him'

2P clitics are typically much freer with regard to the category of potential hosts; frequently they will attach to anything that can occur in first position in a sentence. "First postition" is open to different interpretations in different languages, it could refer to the first word or it could mean after the first constituent. Klavans (1980) cites the following examples from Ngiyambaa, a language in which both interpretations are allowed. The clitic involved here is the second person nominal marker ('=' indicates the clitic boundary):

1.3) a. nadhay=ndu guya dha-yi
tasty =2NOM fish eat-past
'You ate a tasty fish'

b. nadhay guya=ndu dha-yi
tasty fish=2NOM eat-past
'You ate a tasty fish'

Similar situations attain in other languages, such as Serbo-Croatian, as well.

Zwicky (1977) represents one of the first comprehensive clitic typologies, attempting to take into account all of the factors mentioned above: i.e. host preferences, positioning and existence or lack of corresponding free forms. He divides clitics into three distinct groups-- simple clitics, special clitics and bound words-- on the basis of these properties. Zwicky defines a simple clitic as a phonologically reduced version of a free morpheme which becomes subordinate to a neighboring word. These reduced forms occupy the same position in the sentence as their corresponding full forms and so do not exhibit any "special" syntax. To illustrate, Zwicky cites the following example of object pronoun reduction in English:

1.4) a. He sees her
b. [hì síz hr̀] full
c. [hì sìzr]' reduced

The pronunciation in (1.4c) is a casual version of the sentence in (1.4a) and Zwicky notes that simple clitics are usually associated with particular speech styles or speeds.

Special clitics differ from simple clitics in two important ways. First, special clitics occupy a different position in sentence structure than non-clitic elements with the same function. So for example, in the French and Spanish sentences in (1.1) and (1.2), the clitic object pronoun precedes the verb while non-clitic object NP's normally follow the verb

forming a constituent of their own; if instead the verbs were non-finite the clitic would attach to the end of the verb. Second, there is not necessarily a close phonological relationship between a special clitic and any related free form it may have (cf. the Spanish lo veo a el 'I see him', in which el is the free pronominal counterpart to the clitic pronoun lo). Thus, Zwicky concludes that these bound forms are not related to the free forms by phonological rules of any generality.

The third type of clitic in Zwicky's typology, bound words, never have free variants. While bound words attach phonologically to one word they are semantically associated with the entire constituent of which this word is a part. Since it is the constituent as a whole rather than the individual lexical item which is important, bound words can choose from a variety of lexical categories as their host. An example of a bound word would be the English possessive marker 's illustrated below:

1.5) a. The boy's hat
b. The boy who ran's hat
c. The boy who looked up's hat
d. The boy he ran to's hat

In this small number of examples alone the possessive marker is attaching to a noun, a verb, a particle and a preposition though they all are part of an NP constituent.

The problem for this approach is that clitics in many languages do not always fall into these three neat groups. Some clitics may, for example, act like bound words in some respects and like simple clitics in others. Klavans (1980) criticizes Zwicky's typology on just this point also arguing that his approach does not provide a framework in which to describe historical changes in clitic systems or capture similarities and differences between certain clitic types. In particular, Klavans charges that Zwicky's claim concerning the development of bound morphemes--that independent words are reanalyzed as clitics which are then reinterpreted as affixes--lacks motivation in some instances and is historically inaccurate in others. She further objects to the failure of Zwicky's typology to recognize similarities between clitics based on positioning. Klavans cites the example of 2P pronouns in Walpiri and 2P particles in Tagalog: the former are classified as special clitics while the latter are said to be bound words. Thus the fact that clitics seem to be drawn to certain positions in a wide range of languages is obscured.

Klavans rejects earlier typologies of clitics and clitic placement as being too simplistic and suggests that such facts can be given a unified account only by characterizing them in terms of the following five parameters:

P1: Clitic Identity
P2: Domain of Cliticization
P3: Initial/Final
P4: Before/After
P5: Proclitic/Enclitic

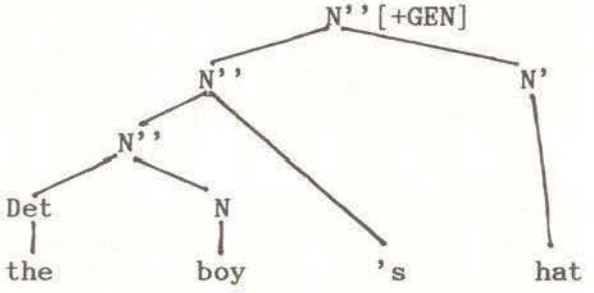
P1 merely refers to a lexical feature [+clitic] by which clitics can be identified by cliticization rules. P2-P4 are concerned with the syntactic placement of the clitic. P2 refers to the node with respect to whose immediate constituents the syntactic position of a clitic is determined. P3 indicates whether it is the first or last immediate

constituent in the domain which is relevant for placement and P4 whether the clitic attaches to the left (Before) or to the right (After) of this constituent. P5 makes explicit the phonological attachment of the clitic; if it attaches to the end of the preceding word it is enclitic, if it attaches to the beginning of the following word it is proclitic. To give an example, the possessive marker in English would have the following values for these five parameters:

P1: English possessive
P2: N''[+GEN]
P3: Initial
P4: After
P5: Enclitic

Since possession is marked on (genitive) NP's this is the domain of cliticization (P2). P3 is initial because the first constituent in NP is marked, i.e. the boy in something like the boy's hat. P4 is after because the marker follows the constituent picked out by P3 and P5 is enclitic because the marker combines phonologically with the preceding material. The tree in (6) illustrates the syntactic positioning:

1.6)




Klavans argues that this typology is superior to Zwicky's because it can capture similarities in syntactic positioning quite straightforwardly and is superior to typologies based soley on a verbal vs 2P distinction in that it allows for a greater range of clitic positions (i.e. eight possible locations per domain). This last feature is, in fact, precisely the problem with her approach; her system is simply too unrestricted. While Klavans claims to have substantiated each of the eight clitic positions (p. 138), the examples she gives are not all from the same domain. As can readily be seen on closer examination it would be impossible to substantiate each of the positions for every domain since some combinations of parameters are nonsensical. Take, for example, the following two parameter combinations:

1.7) a. P1: --
P2: S
P3: Initial
P4: Before
P5: Enclitic

b. P1: --
P2: S
P3: Final
P4: After
P5: Proclitic

Each of the sets in (1.7) would require the clitic to attach to something outside of its own S; neither of these clitic types have been convincingly attested[5].

Of the clitic positions Klavans does attempt to support, some are based on less than persuasive evidence -- a case in point being her categorization of Old Indic Preverbs. Klavans argues for the positioning of Old Indic Preverbs by appealing to the analyisis of Proto-Indo-European preverbs in Anderson (1979). As she herself admits (p. 138) the evidence is only suggestive, it is far from conclusive. Another problem with this typology is that it predicts that every clitic position is just as likely to occur as any other; it gives no explanation of why some positions turn up in language after language in many different families while some positions don't seem to turn up at all. Thus Klavans' analysis is no more informative on this point than Zwicky's and certainly cannot be considered superior to it. Furthermore, Klavens' analysis fails to distinguish in a systematic way between clitics which have free standing counterparts and clitics that don't, overlooking an obvious and, for our purposes, important typological difference. In any case, since our main interest here is not so much with the range of possible positions for clitics (or for that matter with characterizing their historical development), but with their associated properties, Zwicky's system will ultimately be of much more use. Where clitic positioning is relevant we will rely on the standard verbal vs. 2P distinction. In the next section we will take a closer look at the rules of AR and CC and see how reduced auxiliaries and complementizers fit into the framnework assumed above.

1.2 AR and CC.

As we have seen, AR is an optional process by which finite forms of certain auxiliary verbs[6] become dependent on neighboring material. In most dialects reduced auxiliaries show a low degree of selection with regard to the category of the lexical items they attach to. Instead, what seems to be important is the category of the constituent this lexical item is part of[7], as we can see from the examples below:

1.8) a. Pita's a cat.
b. He's a cat.
c. The cat Mary painted red's named Pita.
d. The cat Mary hit's named Pita.
e. The cat Mary talked to's named Pita.
f. The cat Mary fed yesterday's named Pita.

Since reduced auxiliaries have alternate free forms that occupy the same position in the sentence, in Zwicky's typology they would be classified as simple clitics rather than special clitics or bound words. As such we would expect them to display the same type of behavior as other simple clitics and, as we shall see a bit later, this is indeed the case.

While some researchers, most notably Bresnan (1971), have argued that reduced auxiliaries must be treated as proclitic to following material in order to account for sentences in which AR is blocked, most clitic analyses have viewed AR as a rule of enclisis. This, plus the facts that AR applies to finite verb forms and finite verb forms usually follow subject NP's in English, gives reduced auxiliaries something of the appearance of 2P clitics. In fact, this very property is exploited in an interesting discussion of possible causes for dialect variation in sentences involving AR presented in Kaisse (1983b). Notice, however, that

reduced auxiliaries cannot in general be treated as 2P clitics because of the existence of sentences like (1.9b):

1.9) a. John unfortunately is not going.
b. John unfortunately's not going.

In this case the adverb unfortunately is occupying second position and the auxiliary is is reducing onto it. Thus, while it is frequently the case that reduced auxiliaries are 2P, it is not necessarily true. We will return to the question of proclitic versus enclitic treatments in section two.

CC, like AR, is also an optional rule which results in the reduction of a free morpheme (infinitival to) onto preceding material. Since contracted to's, like reduced auxiliaries, have alternate free forms which occupy the same syntactic position, they too would be considered simple clitics in Zwicky's system. Additional examples of CC are given in (1.10):

1.10) a. They wanna be in pictures.
b. They hafta be in pictures.
c. They usta be in pictures.
d. They oughta be in pictures.
e. They gotta be in pictures.
f. They're sposta be in pictures.

In most accounts of CC it would be possible to view contracted to as a verbal clitic since it is either assumed that this reduction can only occur with a few, lexically specified verbs[8] (hence the common name Wanna-Contraction) or with the class of verbs as a whole. The one exception to this is Jacobson (1982) who claims that to can cliticize onto both verbs and adjectives, the particular lexical item involved being irrelevant.

Jacobson bases this claim on sentences with reduced vowels, like those in (1.11)-(1.13), which she says are grammatical for some speakers:

1.11) a. I want tə.
b. He wants tə.
1.12) John seems tə.
1.13) John is expected tə.

She also argues that, even for speakers who disfavor (1.11)-(1.13) there is a sharp contrast between those sentences and ones in which the item preceding the to is not a verb or adjective:

1.14) *I want Sam tə.
1.15) *I persuaded Sam tə.
1.16) *I want very much tə.

All of the sentences in (1.11)-(1.16) seem equally awkward to me, those in (1.14)-(1.16) no more so than the others. But even if there are speakers who share these judgments, Jacobson's conclusions are not warranted. The problem lies in distinguishing actual cliticization from simple phonological vowel reduction. There are two types of evidence in favor of the latter analysis for at least some of Jacobson's examples. One is the critical interplay between her to reduction rule and stress--a known factor in phonological reductions. The second is that reduction of to to

tə is possible even in contexts in which CC is not allowed. An example of this type would be (1.17b):

1.17) a. Who does Pita want tə kiss you?
b. *Who does Pita wanna kiss you?

where (1.17b) is not possible in most dialects. Another example would be Jacobson's sentence (75) given below as (1.18):

1.18) Tə run is no fun.

Since CC involves leftward cliticization the tə in (1.18) could not be a result of the same rule. To account for these sentences Jacobson must posit a second, otherwise unmotivated cliticization rule to perform the same function as well-founded phonological rules. Thus it is clear that Jacobson is attempting to account for too wide a range of phenomena with her rules.

For the purposes of this work I will adopt the view that CC applies to the class of verbs (all verbs and only verbs) with unpredictable phonological effects in some cases and predictable effects in others. The sentences in (1.11)-(1.16) will be attributed to the operation or failure of a phonological reduction rule rather than cliticization. This treatment will allow us to capture the contrast between sentences like (1.11) and (1.12), which some speakers reject, and sentences like (1.19) and (1.20), which they find completely acceptable:

1.19) I wanna.
1.20) He wansta.

Notice, however, that this definition of the domain of CC is not a necessary feature of my analysis. If future evidence persuasively argues in favor of one of the other proposed domains for CC the change can easily be effected using subcategorization. As things stand this slight complication does not seem to be needed. As noted above, this view of CC is consistent with the claim that reduced to is a verbal clitic.

1.2.1 AR, CC and Phonology.

The contrast between (1.17a) and (1.17b) noted above also argues against the claim (suggested by Lakoff (1970), among others) that CC is conditioned by low stress and thus is a phonologically determined rule. Since Jacobson's to reduction rule, which is conditioned by stress, can apply to produce (1.17a) if CC were also stress dependent we would expect it to be able to apply here as well. The fact that CC is ungrammatical in (1.17b) shows that something else is going on in these sentences. A similar argument can be made for AR as well, as pointed out in Kaisse (1983a). While the unstressed auxiliaries in (1.21a) and (1.22a) can be phonologically reduced to [ z] or [ɨz] in most dialects, they cannot be realized as a fully reduced clitic forms, i.e. without any vowel at all, in any dialect:

1.21) a. I wonder how much wine there is in the bottle.
b. *I wonder how much wine there's in the bottle.
1.22) a. John's nicer in the mornings than Harry is at night.
b. *John's nicer in the morning than Harry's at night.

Thus, lack of stress cannot be the determining factor in the operation of AR or CC. The strong correlation that is found between clitics and stresslessness in many languages can be accounted for in other ways; for example, by having rules remove stress from cliticized elements or even by ordering stress assignment rules after cliticization and having them fail to operate on clitics. The point is, we need not and cannot assume that stress is what conditions the operation of the rules under discussion.

In addition to the claim that AR and CC are phonologically conditioned, it has also been suggested that AR and CC are themselves phonological reduction rules. Due to the highly idiosyncratic effects of these rules, however, such an analysis is unworkable as well. As Kaisse (1983a) points out, the phonological rules that would be needed to derive reduced auxiliaries from their full counterparts are either not productive rules of English at all or not productive at all the speech rates which permit AR:

| | | | |
|---|---|---|---|
| 1.23) | a. | is | [s], [z], [ɪz] (or [ɨz]) |
| | b. | are | [ɹ], [r] |
| | c. | am | [m̩], [m] |
| | d. | has | [s], [z], [ɪz] (or [ɨz]) |
| | e. | have | [v], [ɪv] |
| | f. | had | [d], [ɪd] |
| | g. | will | [l̩], [l] |
| | h. | would | [d], [ɪd] |

For example, Kaisse notes that English has no regular rule of [w] deletion, which would be needed in a phonological derivation of (1.23g) or (1.23h). Also, while there are productive rules to delete [h] when it occurs before an unstressed vowel, they apply only in rapid speech. Since AR is possible even at relatively slow speech rates, forms like (1.23d-f) could not be generated. Similarly, even though full vowels can reduce to schwa at all speech rates, the rules which delete schwa entirely are also restricted to fast speech; thus none of the vowelless alternates in (1.23) could be derived at a slower rate either.

Finally, if we examine the reduced alternates of <u>is</u> and <u>has</u> given in (1.23a) and (1.23d) respectively, we notice that they are suspiciously similar to the various allomorphs of the plural, third person singular and possessive morphemes both in form and distribution: only [ɪz]/[ɨz] can occur after stridents while [s] occurs after voiceless non-stridents and [z] after voiced non-stridents[9]. The most general way of accounting for these facts would be to allow the rules which determine the distribution of the allomorphs of these other morphemes to also determine the distribution of the reduced forms of <u>is</u> and <u>has</u>. Since in most recent theories of grammatical organization rules of this type precede phonology proper, the rules which determine when an auxiliary can be realized as its reduced form (as opposed to its full form) must <u>also</u> precede the phonological component and, therefore, must be of a distinct type.

There are similar arguments against treating CC as a phonological reduction, as well. First of all, as with AR, the phonological rules that would be needed to derive CC forms from their full counterparts are not all fully productive. For example, in order to derive the reduced sentence in (1.24b) from its full counterpart in (1.24a):

1.24) a. I want to finish.
b. I wanna finish.

we would need both a degemination rule and a nasal assimilation rule with at least one operating across a word boundary. While rules operating across word boundaries are common in English they are generally restricted to fast or casual speech. Forms like (1.24), however, are perfectly acceptable even in slow, careful styles. Even more disturbing is the fact that there is really no well defined set for these rules to apply to. AR, at least, can be restricted to the class of auxiliary verbs, though not all auxiliary verbs are affected. The verbs which undergo radical phonological changes as a result of CC, however, have no other common properties to set them apart from other verbs. Thus there would be no general way of preventing the derivation of sentences like (1.26b) alongside (1.25b):

1.25) a. I want to live.
b. I wanna live.
1.26) a. I hunt to live.
b. *I hunna live.

In my speech want and hunt differ primarily only in initial consonant, thus there would be no phonological grounds on which to exclude (1.26b). It is obvious, then, that the relationship between want to and wanna needs to be stipulated rather than derived. Since this type of "spelling out" rule is typically found in the morphological component, e.g. take + past tense = took, and the morphological component is typically ordered before phonology, we again have an argument for ordering the rules governing the distribution of full (versus reduced) forms before phonology. Notice that these facts are perfectly consistent with the view that AR and CC belong to a separate component of the grammar reserved for cliticization and ordered between syntax and morphology, as argued for in much of the recent literature (see references, fn. 1).

1.2.2 AR, CC and Morphology.

Another possibility that should be considered here is that reduced sentences are not derived via productive rules at all but, rather, hosts bearing reduced elements are listed separately in the lexicon and assigned the appropriate distribution (e.g. wanna alongside want, John's alongside John). While such an approach to AR is totally unworkable, it is at least plausible in the case of CC. Since the reduced alternates of auxiliaries like is and has appear quite freely with preceding NP's no matter what their composition, it would be impossible to limit the number of different constructions in which they occur. Thus we would either have to list an infinite number of otherwise perfectly regular phrases separately in the lexicon or allow the word that bears the reduced auxiliary, no matter how deeply embedded it may be, to determine the type of matrix VP that is allowed. This is clearly absurd. On the other hand, since reduced to has a much more restricted distribution than reduced auxiliaries, occurring only with verbs that can take an infinitival complement, it would be relatively simple to separately list forms with reduced to and forms without reduced to for each such verb. The forms with reduced to would differ from those without only in that they subcategorize for bare infinitive complements rather than overt infinitive complements.

The problem with this approach, aside from the distributional peculiarities and the redundancy of listing both forms, is that forms like wanna, gotta, etc. do not function syntactically like single words as we would expect if they had separate lexical entries. They do not undergo any type of derivation or inflection, nor are they operated on as a unit by any syntactic rules. In fact, sentences in which they are treated as a unit are judged to be ungrammatical. For example, compare (1.27b) with (1.27c):

1.27) a. John is supposed to drive to Cleveland and Mary is supposed to fly to Toledo.
b. John is sposta drive to Cleveland and Mary is sposta fly to Toledo.
c. *John is sposta drive to Cleveland and Mary is fly to Toledo.

If in fact sposta were a separate lexical item we would expect it to undergo gapping, just like any other verb:

1.28) John will drive from Cleveland to Toledo and Mary will __ from Toledo to Akron.

The fact that (1.27c) is ungrammatical shows that sposta is not a syntactic unit but merely a phonological one. Thus this type of morphological treatment of AR and CC, the lexical approach, cannot work.

A second type of morphological treatment which has been argued for is the view that AR and CC involve affixation rather than cliticization. There are, however, a number of reasons for not believing this to be the case. One such reason is that reduced auxiliaries and contracted to have more properties in common with clitics than they do with affixes. Zwicky and Pullum (1982) present the following criteria for distinguishing between simple clitics and affixes (Z&P. p.3):

1.29) a. Clitics exhibit a low degree of selection with respect to their hosts, while affixes exhibit a high degree of selection with respect to their stems.
b. Arbitrary gaps in the set of combinations are more characteristic of affixed words than of clitic groups.
c. Morphophonological idiosyncracies are more characteristic of affixed words than of clitic groups.
d. Semantic idiosyncracies are also more characteristic of affixed words than of clitic groups.

If we measure the results of AR against the principles in (1.29) we see, as Zwicky and Pullum themselves point out, that reduced auxiliaries are almost a paradigm example of simple clitics.

Examining just the sentences given in (1.8) above we find examples of an auxiliary verb reducing onto a noun, a pronoun, an adjective, a verb, a partical, and an adverb. From this we can see that, though there may be general restrictions on the preceding constituent in some dialects, the category membership of the word the auxiliary actually attaches to is not important; reduced auxiliaries do indeed exhibit a low degree of selection. Furthermore, unlike affixes, there are no cases in which a particular lexical item idiosyncratically blocks the application of AR.

There are cases in which AR is disfavored (not blocked) with particular lexical items, but these are for perfectly straightforward phonological reasons. The phonological effects of combining a reduced auxiliary with its host are also perfectly straightforward. While irregular plural or past tense forms are quite common, the phonological variations in reduced auxiliaries are fixed and predictable from the phonological and morphological properties of the host. Finally, there are no cases in which the semantic contribution of the reduced auxiliaries is in any way different from the semantic contribution of the corresponding full form.

Though contracted to's do not fare quite so well with respect to the criteria in (1.29) they do, nonetheless, have some distinctly non-affixal properties. The fact that CC does not allow a wide range of categories to act as host does not necessarily reflect on its status as a cliticization rule since, as we saw above, a large number of clitics are restricted to verbal hosts. This is just one way in which contracted to's are less like simple clitics than reduced auxiliaries are. Since, by our definition, CC will reduce to onto any verb we do not have arbitrary gaps in the set of possible combinations. We do, however, have morphophonological idiosyncracies in a few of these combinations. Notice though that the total number of such idiosyncracies is much lower than for verbal paradigms. Notice also that such irregularities can occasionally be found in known clitic groups as well (Spanish le lo -> se lo), they are merely less frequent. As with reduced auxiliaries, the semantics of contracted to is entirely compositional. In sum then, contracted to does not exhibit any behavior that cannot be attributed to some type of clitic (though not always simple clitics) though it does lack certain properties frequently found in affixes.

There are other reasons for rejecting an affixal analysis of AR and CC as well. For one thing, treating these rules as affixation would greatly complicate the morphology of English. In addition to paradigms like (1.30):

1.30) a. I want
b. you want
c. he, she, it wants
etc.

we would have ones like the following:

1.31) a. I wanna
b. you wanna
c. he, she, it wansta
etc.

This would be true for every verb that underwent CC (i.e. for every verb in the language that takes an infinitival complement). We would also have to somehow insure that such verb forms are followed by verb phrases beginning with bare infinitives. This would be a novel situation in that it would be the affix subcategorizing the following material rather than the verb itself. The situation with AR would be even worse since reduced auxiliaries can attach to elements from so many different categories; we would in effect be creating a group of affixes that can attach to almost any word in the language but are semantically associated with the entire sentence. Again, this is clearly absurd. Consider also the fact that

affixed words can be treated as units by syntactic rules unlike the products of AR and CC, as we saw above. In the case of AR, as Zwicky and Pullum point out, such a syntactic rule would be almost inconceivable.

Perhaps the most persuasive reason for rejecting an affixation analysis of AR and CC, however, is the fact that both operations are sensitive to aspects of the sentence other than just the word they are attaching to. Compare the following pairs:

1.32) a. Who does Pita wanna see?
b. *Who does Pita wanna see you?
1.33) a. Who's going?
b. *Who's?

While all the sentences in (1.32) and (1.33) are grammatical with their corresponding full forms, only the (a) sentences allow reduction. This is an important difference between affixation rules and AR and CC; while the conditions governing the combining of affixes with their stems are purely morphological and lexical, those governing the application of AR and CC seem to be syntactic in nature. This argues in favor of a separate, non-affixal analysis of AR and CC. Thus, all things considered, the clitic analysis of reduced auxiliaries and contracted to is more strongly supported by the evidence and we can conclude that AR and CC are, in fact, rules of cliticization rather than affixation or phonological reduction.

1.3 Cliticization and Syntax.

In the preceding sections it was argued that AR and CC are conditioned by syntactic structure rather than by phonological, morphological or lexical considerations. It should be noted that this is very different from the claim that AR and CC are themselves syntactic rules. In fact, contra Bresnan (1971), there does not seem to be very much evidence for the claim that cliticization rules belong in the syntactic component of the grammar. Notice, first of all, that there are no syntactic rules whose operation depends on the application of a cliticization rule. Nor, as we will see in section two, are there any syntactic operations that are bled by a cliticization rule either. Furthermore, cliticization rules are of a very different type than other syntactic rules dealing, as they do, with units smaller than words rather than entire words and phrases. This is all consistent with the view that rules like AR and CC form their own component in the grammar, one dealing with the production of phonological words rather than syntactic words. While this is a much more restricted model of grammar in that it severely limits the range of possible rule interactions, it is in no way predicted by current transformational frameworks. In section three I will show that, given a GPSG syntax, this type of organization falls out automatically; thus supporting a conclusion reached on independent grounds by many others (see references fn. 1).

1.4 Conclusions.

In the preceding discussion I have argued for the claims that 1) AR and CC are, in fact, synchronic rules of grammar, 2) that they are best analyzed as belonging to a separate component of the grammar reserved for rules of that type and 3) that the primary factor in determining the applicability of AR or CC is the syntactic structure of the candidate sentence. In the following two sections I will discuss the issues of how these conditions on syntactic structure should be formulated and what the optimal analysis shows about the grammar as a whole.

## 2. Previous Analyses

A number of different analyses of CC and AR have been proposed over the years with widely different views of how the cliticization process fits into the framework of a grammar. Most, if not all, of these analyses have recognized the need to refer to syntactic structure when describing the conditions under which these rules apply. These treatments can be loosely grouped into three types: those requiring some sort of explicit global reference, those involving the transformational cycle, and those appealing to some form of trace element. In what follows I will briefly review some of the more influential of these past analyses while pointing out some of the problems these treatments have had. I will return to the discussion of the place of cliticization rules in the grammar in section 3.

### 2.1 Global Rules.

Perhaps the best known discussion of AR and CC is the "global rule" analysis given in Lakoff (1970). Lakoff sees both AR and CC as purely phonological reductions and argues that since they are sensitive to aspects of syntactic structure they must be global rules. Lakoff bases his formulation of the conditions on AR on facts about where be can reduce first noticed by King (1970):

2.1) a. i. There's a man in the room.
ii. *I asked which men there're in the room.
b. i. It's hot.
ii. *. . . and hot it's.
c. i. You said the concert's in which park?
ii. Which park did you say the concert's in?
iii. *In which park did you say the concert's?
d. i. Kim is to leave and Sandy's to, also.
ii. *Kim is to leave and Sandy's, also.

While sentences like (ai, bi, ci, cii, and di) allow be to contract, the corresponding sentences in which Wh-Movement, Topicalization, or VP-Deletion have disturbed the complement of the auxiliary do not. Lakoff (p. 631) cites the following generalization "If there is a constituent immediately following be, and if by any transformation that constituent is deleted, then the be cannot contract." The problems with this formulation are well known. While Lakoff can account for the contrasts in (2.1), his analysis makes incorrect predictions about the grammaticality of the sentences in (2.2) and (2.3):

2.2) John's to force himself to stop.
2.3) a. Where's the library?
b. What's a global rule?
c. How fat's your cat?
d. In which city's the conference?

In his transformational framework, no matter how you order to Insertion and Equi NP Deletion the be in (2.2) would be followed by a movement or deletion site and, therefore, should not be contractable. Similarly, the sentences in (2.3) would be ruled out since they involve not only the

movement of the constituent following be, but the movement of be itself.

Lakoff also proposes a global constraint on CC to account for the contrast in meaning between sentence pairs like the following, first noticed by Horn (cited in Lakoff (1970)):

2.4) a. Teddy, I want to succeed.
b. Teddy, I wanna succeed.

Sentence (2.4a) is ambiguious between the readings I want Teddy to succeed and I want to succeed Teddy while sentence (2.4b) can only have the second interpretation. Lakoff concludes from such sentences that CC is blocked if at any stage in the derivation an NP had intervened between the verb and to. The deep structures for the sentence in (2.4a) would presumably be those in (2.5):

2.5) a. I want [Teddy succeed]
b. I want [I succeed Teddy]

Notice that for this analysis to work Lakoff must explicitly order to Insertion after Equi NP Deletion (to permit contraction in (2.5b)) but before the rules responsible for topicalization (to block contraction in (2.5a)). This seems to be the only motivation for such an ordering.

Another problem with this analysis is that it fails to block contraction in sentences like the following, taken from Pullum and Postal(1982):

2.6) a. To regret what one does not have seems like to want.
b. ?It seems like to want to regret what one does not have.
c. *It seems like to wanna regret what one does not have.
2.7) a. I don't want anyone [who continues to want] to stop wanting.
b. *I don't want anyone [who continues to wan]na stop wanting.
2.8) a. I want to dance and to sing.
b. *I wanna dance and to sing.
2.9) a. I don't need or want to hear about it.
b. *I don't need or wanna hear about it.

Though these sentences satisfy the condition on intervening NP's, none of them allow contraction.

A revised version of Lakoff's constraint on AR is presented in Kaisse (1983a) where it is suggested (p.93) that the original condition be interpreted as in (2.10):

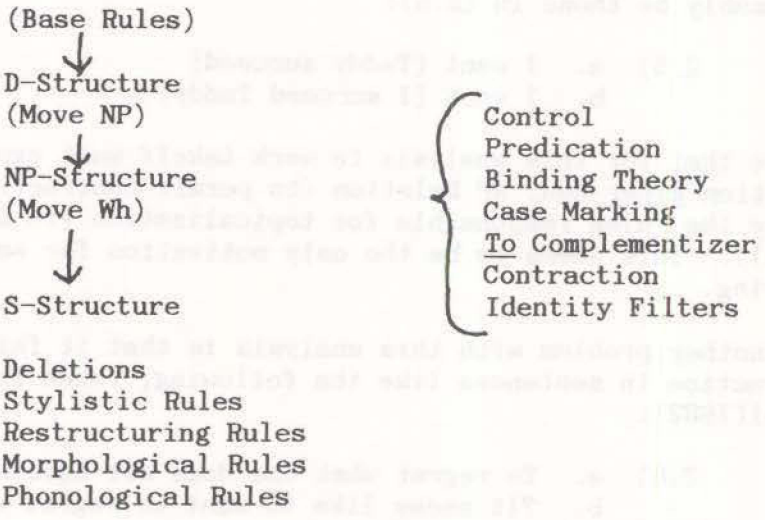
2.10) Auxiliary Reduction may not apply if the element following the auxiliary is not the same as the element that follows it at the stage in the derivation prior to all movements and deletions.

In addition to the sentences in (2.1), this formulation accounts for the ungrammaticality of AR in something like (2.11b) in which an element has been inserted following the auxiliary:

2.11) a. He is, I should think, a bit tired.
b. *He's, I should think, a bit tired.

The constraint in (2.10) does not, however, make the necessary distinction between rules like Equi and There Insertion (which do not block reduction) on the one hand and Wh-Movement and Topicalization (which do block AR) on the other. Nor does it sufficiently limit the class of hosts to those permitted in Kaisse's dialect. To remedy these defects Kaisse adopts a modified "split T" model of grammar in which "move NP" rules are distinguished from "move Wh" rules[10]. AR is then made sensitive to the level of structure resulting from the "move NP" rules and the set of possible hosts to AR restricted to NP's. The model of grammar Kaisse assumes is shown in (2.12):

2.12)




Following Pullum and Zwicky (forthcoming) cliticization rues are treated as part of a separate component, labeled here as "Restructuring Rules". Given this model, Kaisse's restriction on AR is as follows:

2.13) X NP AUX Y Z -----> 1, 2#3, Ø, 4, 5
1 2 3 4 5
where 2 c-commands 3, and 4 follows 3 at NP-Structure

This says that AR is possible just in case the host is a noun phrase which c-commands the auxiliary and the element following the verb to be cliticized followed it at NP-Structure. Thus, the starred sentences in (2.1) are blocked since rules have applied to the NP-Structure which have altered the material following the verb. Presumably a sentence like (2.2) would be generated without a subject NP in the lower clause, thus contraction is possible. As it stands the rule in (2.13) also incorrectly predicts that sentences like (2.3a) and (2.3b), repeated below, are ungrammatical:

2.3) a. Where's the library?
b. What's a global rule?

To account for cases like these Kaisse includes a rider on her restriction which allows reduction onto a nonlexical item, such as a Wh-word, as long as the element following the verb meets certain criteria. This rider is given in (2.14):

2.14) In addition, if 2 is a (monosyllabic) pro-form it need not be an NP, and it suffices that X [Y: AB] not mark a movement or deletion site.

Kaisse's constraint on AR makes many other predictions as well. Thus all of the sentences in (2.15)-(2.20) will also be blocked by (2.13):

2.15) Which dog's he buying?
2.16) Not only's Louis smart, he's also a varsity rower.
2.17) On which day's John leaving?
2.18) a. Speaking tonight's a famous reporter.
b. Speaking tonight's been a famous reporter.
2.19) a. More important's her insistence on honesty.
b. More important's been her insistence on honesty.
2.20) Under this slab's buried Joan of Arc.

(examples based on Kaisse (1983a)). Sentence (2.15) is bad because, due to the application of Subject Auxiliary Inversion (SAI), the element following be at the time cliticization operates is not the same as the element which followed it at NP-Structure. Sentences (2.16) and (2.17) are rejected on two counts: SAI has applied in these sentences (triggered by various preposing rules) and the host for the clitic is not an NP. While the sentences in (2.18), (2.19) and (2.20) do not involve SAI (cf (2.18b) and (2.19b)) they still fail the NP host condition.

While the constraint in (2.13) may adequately describe Kaisse's dialect, it does so at the cost of employing an extremely powerful mechanism -- a global rule. In addition to this, there are dialects in which all of the sentences given in (2.15)-(2.20) are perfectly grammatical[11]. This poses a particularly difficult problem for Kaisse's analysis since some of these sentences violate both conditions of her constraint at the same time. Thus there would be no way of generalizing Kaisse's constraint to include this other dialect. Since an analysis which accounts for different dialects with a related set of rules is to be preferred over one which treats them with entirely separate rules, Kaisse's constraint is less than satisfactory.

Kaisse (1983b) presents a modified version of this analysis in which the condition on preceding context is altered to bring AR more in line with the behavior of similar clitics in other languages. Kaisse argues that reduced auxiliaries are second position clitics and, as such, should not be sensitive to the category membership of their host. Thus she replaces her NP host condition with the following:

2.21) An auxiliary may only cliticize onto the first word of its S.

This constraint rules out the sentences in (2.15)-(2.20) since the various preposing operations involved -- Wh-Movement, Comparative Preposing, PP Fronting, etc. -- move material into COMP and outside the

domain of S. As a result, the auxiliary itself is the first element in the S in these sentences and, thus, cannot appear in reduced form. However, according to Kaisse's own article sentences like (2.22) involve Wh-Movement of the subject phrase:

2.22) Which man's going to win?

Since Wh-Movement would insert the subject phrase into the COMP node, Kaisse's analysis predicts that sentence (2.22) is also ungrammatical and for the same reason as (2.15) and (2.17). This judgment is not confirmed in any dialect studied to date including Kaisse's. Kaisse notes a similar problem with sentences like:

2.23) Jack is the man who I bet's going to win.

Since the auxiliary would be the first element in the embedded S, the structure should not permit reduction[12].

The constraint in (2.21) does, however, allow for relatively more dialect variation than the NP-host condition since the domain involved can be easily modified. Thus Kaisse can account for the fact that sentences like (2.18b), (2.19b), and (2.20) with preposed elements in COMP are perfectly grammatical for many speakers by changing the S in (2.21) to S' for these dialects. However, if Kaisse is still assuming the restriction on following context given in (2.13)[13], she cannot explain the grammaticality of the parallel sentences in (2.18a) and (2.19a) or the sentences in (2.15), (2.16) and (2.17) for these same speakers. Not only does the revised constraint in (2.21) incorrectly predict the facts of Kaisse's own dialect, it still fails to account for the judgments found in other dialects. A GPSG analysis, on the other hand, presents a unified analysis of both.

2.2 Cyclic Treatment.

Bresnan (1971) proposes making rules of cliticization such as CC and AR part of the transformational cycle rather than including them in the phonology as Lakoff did. In her analysis of CC, to can cliticize leftward onto the proper type of verb if they are adjacent during that verb's cycle. Thus sentences like (2.24a) will be allowed to undergo CC since the subject of the lower clause is removed by Equi on the want cycle, leaving it adjacent to to.

2.24) a. You want [you kiss who]
b. Who do you wanna kiss?

A sentence like (2.25b), however, will not be produced since who is moved to the front of the sentence from its position between want and to by Wh-Movement on the higher, S', cycle. Hence, since CC is presumed to be cyclic, it never gets a chance to apply:

2.25) a. You want [who kiss you]
b. *Who do you wanna kiss you?

The ambiguity contrast found in (2.4a) and (2.4b) would be parallel to this example: only one of the readings of (2.4a), that corresponding to the deep structure in (2.5b), has want and to adjacent on the want cycle; the other has an intervening NP. Notice that this approach also accounts

for the ungrammaticality of (2.6c) and (2.7b). Assuming strict cyclicity, the to in these sentences would not be eligible for contraction onto the want since they are not part of want's complement. Crucially, however, it does not explain the failure of CC to apply in sentences like (2.8b) and (2.9b). In these cases the to phrase is the complement of an appropriate verb and the two are adjacent on that verb's cycle, and yet contraction does not take place.

Bresnan's cyclic analysis of AR also runs into problems. In an attempt to explain why the material following the verb should be relevant to reduction, Bresnan reanalyzes AR as a rule of procliticization in which the auxiliary is attached to the front of the next word. In this treatment the starred sentences in (2.1) would have to have structures such as the following:

2.26) a. *I asked which men there 're_ in the room.
b. *. . . and hot it 's_.
c. *In which park did you say the concert 's_?
d. *Kim is to leave and Sandy 's_, also.

These could not, however, be generated once cliticization has taken place, since the transformations involved are not defined over subparts of words.

This type of analysis fails on both syntactic and phonological grounds. As Lakoff (1972) points out, a proclitic treatment of reduced auxiliaries would be very peculiar given the fact that clitic has and is--like plurals, possessives, past tense and third person singular markers--assimilate in voicing to what precedes not what follows. In order to account for this fact Bresnan would have to posit an otherwise unmotivated word external process to perform precisely the same function as a well documented word internal process, thus missing an obvious generalization and unnecessarily complicating the phonology.

A syntactic argument against this analysis is given in Wood (1979). Wood notes that Bresnan's treatment of AR cannot account for the grammaticality of sentences like:

2.27) Herb's going and Jerome is __ too.

If AR is cyclic then it applies or fails to apply on the same cycle for each conjunct. Therefore, after the first cycle the lowest verb phrase of the left conjunct would be 'sgoing while the lowest verb phrase of the right conjunct would be going. Thus the identity condition on VP Deletion would not be met and the sentence in (2.27) could not be generated.

A final problem with this approach stems from the nature of cliticization rules in general. Clitic elements, unlike some types of affixes, do not change the category of their hosts; e.g. wh-words with clitics attached are the same category as they would be without the clitic. Since Bresnan views cliticization as a process by which elements become syntactic dependents of preceding or following elements, in order to prevent sentences like:

2.28) a. *I asked 're-which men there __ in the room.
b. *'s-In which park did you say the concert __?
c. *. . . and 's-hot it.

one would have to somehow build into each movement rule a clause that examines the structure of the constituent in order to detect the presence of any clitic elements and prevents the rule from applying if such an element is found. This would complicate these rules enormously. Furthermore, as we saw in section one, Bresnan's basic assumption--that cliticization rules can and should be included in the syntactic component of the grammar--is not supported by rule interaction facts.

2.3 Trace Theories.

Perhaps the most frequently appealed to type of analysis is one involving some sort of trace element. In such analyses, cliticization is possible only if traces do not appear in the relevant positions in syntactic structure. What form these traces take and how, precisely, they arise is a matter of considerable variation from theory to theory. Selkirk (1972) proposes an analysis in which traces take the form of extra word boundary markers which serve to block the destressing rules that feed various cliticizations. According to her analysis, word boundary symbols flank members of major categories in deep structure. When transformational rules move or delete elements they leave the position of these boundaries unaffected. When a moved item is adjoined elsewhere in the sentence new boundary markers are created. Selkirk also includes a convention by which redundant internal boundary symbols are deleted in the configurations W#]#]Z and Z[#[#W as long as the outermost bracket is not labeled S'. The destressing rule relevant to our concerns is Selkirk's "Monosyllabic Rule" which removes stress from monosyllabic dependents that are followed by at most one word boundary symbol followed by a word with a stressed vowel. Thus a sentence like (2.laii) could not be generated since after Wh-Movement the (simplified) structure would be as in (2.29):

2.29) $[_S$I asked $[_S$'#$[_{COMP}$[#which men#]]$[_S$#$[_{NP}$there#] $[_{VP}$#are [##]$[_{PP}$#in the room#]]]]]

The auxiliary are in (2.29) is followed by a series of two word boundary symbols and, therefore, cannot undergo the Monosyllabic Rule. As a result, the stress on are is not reduced and it cannot undergo cliticization.

There are a number of problems with this analysis[14] one of which hinges on the very feature which allows ungrammatical sentences like (2.laii) to be excluded. If boundary markers are left behind by all movement and deletion rules, then a sentence like (2.30a) with a deep structure as in (2.30b) would incorrectly be blocked from undergoing destressing and subsequent cliticization as can be seen from the surface structure in (2.30c):

2.30) a. John is to leave as soon as possible.
b. $[_S$John is $[_S$'#COMP$[_S$#$[_{NP}$#John#]$[_{VP}$#leave. . .
c. $[_S$John is $[_S$'#COMP$[_S$#$[_{NP}\emptyset$#]$[_{VP}$#to#leave. . .
d. John's to leave as soon as possible.

After Equi applies to the lower S there will be a series of four boundary markers, one of which will be removed by the redundant boundary symbol convention discussed above. Since there are three boundary symbols between is and the nearest following word with a stressed vowel

the conditions for the Monosyllabic Rule are not met. Since is cannot be destressed it also cannot cliticize, thus the sentence in (2.30d) cannot be produced.

As Postal and Pullum (1978) point out, Selkirk's analysis fails (in precisely the same way) for CC as well. Thus a sentence like (2.31a) would have a surface structure as in (2.31b):

2.31) a. I want to go.
b. $[_S$I want$[_{S'}$#COMP$[_S$#$[_{NP}\emptyset$#$][_{VP}$#to go#$]$

Since there are extra boundaries between want and to, destressing and cliticization are incorrectly blocked.

The other types of trace theories proposed thus far have similar problems. Those put forth in Chomsky (1976,1977) assume that movement transformations leave traces in surface structure to mark the position of an element before the rule applied. Postal and Pullum (1978) argue, however, that these theories are incompatible with Chomsky's claim that Wh-Movement is successive cyclic since traces will be overgenerated in COMP position. Thus a sentence like (2.24b) would have the (pre-contraction) surface structure in (2.32):

2.32) [[who do you want[[t] to kiss t]

which is not compatible with CC. Chomsky and Lasnik (1977) attempt to correct this prediction by proposing a rule to freely delete material in COMP positions. However, as Postal and Pullum note, since the COMP node itself is not pruned by this rule want and to are still not structurally adjacent and, therefore, cannot cliticize.

Chomsky (1980) deals with this problem in another way. He argues that traces left by Wh-Movement in non-COMP positions in the clause are case-marked traces and count as syntactic material whereas traces in other positions do not. Since these case-marked traces count as syntactic material, they block contraction. By including this abstract feature, Chomsky is able to distinguish between unbounded dependencies (which do not allow contraction across a t) on the one hand, and Raising and Equi constructions (which do allow this contraction) on the other. Also accounted for is the possibility of cliticization in sentences like (2.32); since the t intervening between want and to is in COMP position it is not a case-marked trace and does not block CC.

There are, however, some problems with these claims. Pullum and Postal (1982) argue that Chomsky's assumptions make it impossible for any dialect of English not to have case marked traces and thus does not account for "liberal" dialects which accept cliticization in sentences in which a marked trace should intervene between the verb and to. In such a dialect sentences like (2.25b) are perfectly fine:

2.25) b. Who do you wanna kiss you?

Furthermore, they point out that since none of the examples in (2.6)-(2.9), repeated below, involve the intervention of a case-marked trace between the want and to, Chomsky's theory fails to account for why cliticization is blocked in each case:

2.6) a. To regret what one does not have seems like to want.
b. ?It seems like to want to regret what one does not have.
c. *It seems like to wanna regret what one does not have.

2.7) a. I don't want anyone [who continues to want] to stop wanting.
b. *I don't want anyone [who continues to wan]na stop wanting.

2.8) a. I want to dance and to sing.
b. *I wanna dance and to sing.

2.9) a. I don't need or want to hear about it.
b. *I don't need or wanna hear about it.

Obviously in these cases mere reference to the position of case marked traces is not enough; one must also take into account other aspects of clause structure, something that Chomsky does not do. Pullum and Postal themselves argue that the underlying failure of trace theories stems from "an unwarranted and unjustified assumption made at the outset and apparently never questioned by TT [trace theory] advocates. This is that linear contiguity is fundamental to the description of contraction" (p. 130). They, however, claim that adjacency is <u>not</u> the primary prerequisite to contraction and propose the following "relational generalization":

2.33) A contraction trigger V can have a contracted form with infinitival <u>to</u> only if:
a. <u>to</u> is the main verb of the initial direct object complement of the matrix clause whose main verb is V;
b. the final subject of the complement is identical to the final subject of the matrix.

If, however, adjacency is not a primary prerequisite we would expect sentences such as (2.34a) to allow contraction since it satisfies both of the conditions specified in (2.33). As we can see from (2.34b), CC is not acceptable here[15]:

2.34) a. I want very much to finish this chapter.
b. *I wanna very much finish this chapter.

From this we must conclude that conventional wisdom is correct after all and linear contiguity is in fact a necessary part of the condition on CC.

## 2.4 <u>Towards a GPSG Approach</u>.

The treatment of AR and CC I am going to argue for here is, more or less, a trace analysis too, albeit one that refers to clause structure as well. The difference between my analysis and other such analyses is that in a Generalized Phrase Structure grammar different predictions are made about which syntactic structures contain traces or gaps. Thus a GPSG analysis avoids the problem found in transformational treatments of how to distinguish the movement and deletion rules which block cliticization from those that don't. In the next section I will briefly outline the basic tenets of GPSG and show how they can lead to a simple and elegant statement of the conditions governing AR and CC in the dialects discussed here.

## 3. A GPSG Analysis of the Data

One difficulty in presenting a unified account of AR and CC within GPSG is that the framework itself has been through a number of extensive revisions in a relatively short amount of time. In the following section I will briefly summarize the most recent version of GPSG as presented in Sag and Klein (1982) and Gazdar and Pullum (1982). Throughout this chapter I will attempt to standardize the varying notation as much as possible while maintaining the basic content of the rules; though I adopt the familiar S/NP/VP symbols whenever possible for perspicuity, it should be remembered that GPSG embraces an X-bar philosophy. When necessary I will use the symbol "a" to stand for the Greek letter alpha and "b" for the letter beta.

### 3.1 The Framework.

GPSG is a surfacy theory of generative grammar in which structural descriptions are assigned to sentences solely on the basis of phrase structure rules; no use is made of transformations or coindexing devices and only one level of structure is defined. The set of immediate dominance (ID) rules are the syntactic basis of a GPS grammar. ID rules have the form:

$$\langle n;\ A \rightarrow B,\ C,\ D \rangle$$

where B, C and D are the categories that A dominates and n is a rule number which acts as a subcategorization feature on any lexical items introduced by the rule. The relative order of B, C and D is given by the set of linear precedence (LP) statements. An example of a LP rule of English would be:

$$NP > PP$$

This says that in any ID rule which introduces both an NP and a PP, the NP will always occur before the PP. In order for a PS rule to be included in the grammar it must be consistent with at least one ID statement and with every LP rule.

Perhaps the most intriguing aspect of GPSG is its use of a metagrammar to capture the generalizations that hold between ID rules and govern their operation but which are not expressed directly with the ID-LP statements. The metagrammar uses two types of devices to capture these generalizations: a set of metarules and a set of rule extension principles. Metarules are a means of expanding the set of ID rules in a rule-governed way; that is, they map ID rules into new ID rules. Metarules have the general form indicated below:

3.1) $a \rightarrow b_1, \ldots, b_n$

$\downarrow$

$a' \rightarrow b'_1, \ldots, b'_n$

This is interpreted as saying that if the ID rule $a \rightarrow b_1, \ldots, b_n$ is in the grammar then the ID rule $a' \rightarrow b'_1, \ldots, b_n$ will also be in the grammar. Since by convention rule numbers are preserved under metarule application they are not specifically mentioned in (3.1)[16].

Rule extension principles "flesh out" these schematic ID rules into fully specified PS rules complete with semantic interpretations. These principles are of two types: rule translation principles and feature instantiation principles. The rule translation principles predict the form of semantic translation rules on the basis of the ID rules and the semantic types assigned to the categories they contain. They thus provide a mapping from ID rule doubles, consisting of a rule number and ID rule, into ID rule triples, which contain in addition Montague-like translation formulae.

Features play a very important role in the GPSG framework. In fact, in the most recent versions of the theory much of the work previously done by metarule is now handled by the feature system and the rules which govern feature assignment (i.e. the feature instantiation principles). Not surprisingly the feature system in GPSG has become quite complex[17]. As is also the case with current versions of transformational grammar, syntactic categories in GPSG are not seen as simple unanalyzable node lables but are instead assigned an internal structure consisting of features. The major innovation in the GPSG system is the idea that these features may take other features as their coefficients[18]. Thus the structure of features is defined as follows:

> 3.2) A feature consists of a feature name optionally followed by one or more features or feature names. Features begin with a left bracket and end with a right bracket. (Gazdar and Pullum (1982), p.3)

Syntactic categories are simply a type of feature, in particular one whose feature name is CAT or CAT'. The internal structure of CAT and CAT' is given below:

3.3) a. [CAT' CAT FOOT]
b. [CAT BAR HEAD]

The feature BAR indicates the phrasal level of the category in an X-bar syntax; it takes as its coefficient a number from 1 to 3 or the feature LEXICAL. For purposes of subcategorization, rule numbers are assigned as the value of the feature LEXICAL. The feature HEAD consists of the syntactic information that is shared between phrases and their heads. This information is divided between the features MAJOR and MINOR as shown in (3.4):

3.4) a. [HEAD MAJOR MINOR]
b. [MAJOR {+N,-N} {+V,-V}]
c. [MINOR AGR CASE . . .]

The feature FOOT contains information about other types of syntactic dependencies that hold between phrases[19]. The internal structure of FOOT and its coefficients is shown in (3.5):

3.5) a. [FOOT SLASH WH REFL]
b. [SLASH CAT]
c. [WH AGR WHMOR]
d. [REFL AGR]

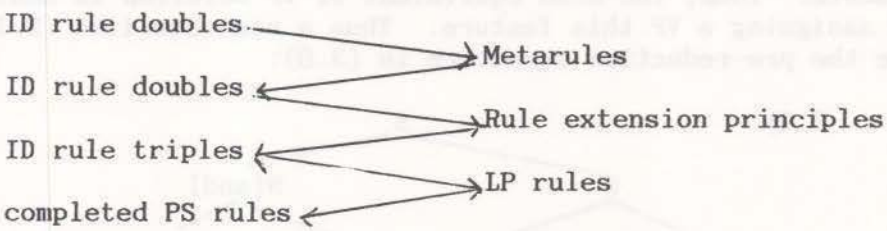
The FOOT feature SLASH is used to encode information about gaps in unbounded dependencies, it takes as its value a category. WH is used in the treatment of Wh expressions; it takes two other features as its coefficients, an agreement feature (AGR) and a feature to encode the morphological type of the Wh word (WHMOR). The feature REFL marks reflexive expressions and also takes AGR as its value. We will have more to say about FOOT features later.

Feature instantiation principles are responsible for ensuring the proper distribution of features in rules. They can be thought of (Sag and Klein (1982), p. 97) as "axioms that must be satisfied by an IDR triple if it is to be an instantiated extension of a given IDR double". Feature values can be assigned in a number of ways: they can be specifically mentioned in an ID rule or metarule, they can be freely assigned in accordance with any default values an item may have (for example, an NP in English is [-CASE], i.e. accusative, unless otherwise specified), or they can be set equal to some other set of features by virtue of special conventions. The special conventions we will be most interested in here are the Head Feature Convention (HFC) and the Foot Feature Principle (FFP).

To put it very simply, the HFC requires the coefficients of HEAD in the mother category and the head daughter to be the same. The "head daughter" is identified on the basis of syntactic category and bar level. For example, given a phrase X'' the head daughter will be either an X'', an X' or an X that it immediately dominates. If X'' dominates more than one of these then the one with the fewest bars will be the head; if it dominates none of these then X'' will have no head[20].

The FFP is responsible for the distribution of FOOT features. Again, very simply put, the FFP says that any FOOT features not assigned to daughters by specific rules must also appear on the mother node. There is nothing to prevent more than one daughter from carrying the same value for a FOOT feature or from carrying different FOOT features altogether, though they are blocked from having different values for the same FOOT feature since there would be no way to encode this on the mother node. Thus, for example, a VP cannot simultaneously have both an NP gap and a PP gap since SLASH can have only one value for CAT.

Rule translation and feature instantiation are two aspects of the mapping from ID rule doubles to ID rule triples. Sag and Klein (p. 98) point out that since both can affect how constituents are linearized in a language the set of rule extension principles must operate before the LP statements. Their view of how the grammar is organized is given below:




3.2 Auxiliaries in GPSG.

My approach to auxiliaries is basically the same as that presented in Gazdar, Pullum and Sag (1981) with one small exception having to do with the treatment of the copula. In that work a verb that is [COP, AUX] can take any of the following complements:

3.6) VP[PRP] is going
VP[PAS] is given
VP[INF] is to leave
VP[PRD]

where a VP[PRD] "merely consists of a predicational X" [i.e. XP]" (GPS, p. 9). I will simplify this somewhat and say that any [COP] verb that takes an XP[PRD] as its complement is also an [AUX], where an XP[PRD] can be any of the following: AP[PRD], VP[PRD, PAS], VP[PRD, PRP], VP[PRD, INF], NP[PRD], PP[PRD]. This results in slightly different tree structures (i.e. no dominating VP node for AP's, NP's, and PP's) and is more in keeping with more recent GPSG works. Notice that by this definition the verb in (3.7a) is not an auxiliary, since its complement is an S (i.e. V''') rather than an XP, and therefore does not undergo AR, as we can see from (3.7b):

3.7) a. The fact is Pita left.
b. *The fact's Pita left.

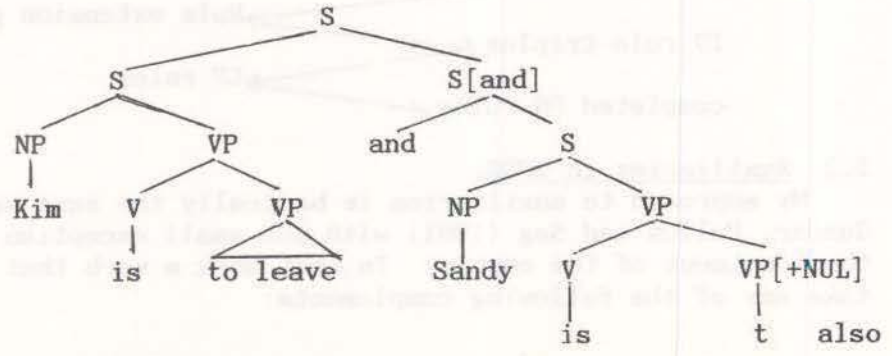
Thus, sentences like (3.7b) will not be considered in our later discussion of the conditions under which AR takes place. Again, where necessary I will modify rule notation to be consistent with this treatment of the copula and its complements.

## 3.3 The Distribution of Traces in GPSG.

As we saw in chapter 2, the problem with transformational analyses of AR and CC is that they fail to distinguish in a general way between operations that block contraction-- such as Topicalization, VP Deletion and most forms of Wh-Movement--and those that don't--i.e. Equi and There Insertion. A GPSG analysis of the same data does not run into this problem because in GPSG there is a natural distinction between the two sets of constructions: the GPSG equivalents of the former involve the introduction a phonologically null element while the GPSG equivalents of the latter do not. Thus the distribution of these null elements can be used to state the conditions governing the application of AR and CC.

Since GPSG is a non-transformational monostratal theory, null elements do not arise through the operation of movement or deletion rules. Instead, the distribution of traces is governed by the feature system and metarules. Categories which are marked with the feature [+NUL] do not receive a phonological representation and are, therefore, trace elements. Thus, the GPSG equivalent of VP Deletion is achieved simply by assigning a VP this feature. Thus a sentence like (2.1dii) would have the pre-reduction structure in (3.8):

3.8)



where t is an abbreviation for VP[+NUL][21]. Since trace elements retain their other category features, null categories will be linearized by the LP statemants just like their non-null counterparts.

In my analysis, traces are also introduced by a version of Slash Termination Metarule 1 (STM1), one of the rules used in Gazdar, Klein, Pullum and Sag (1982) (GKPS henceforth) to "eliminate" unbounded dependencies. This version of STM1 is given below[22]:

3.9) STM 1

$\underline{a}$ -> W, $\underline{b}$[-CASE]

$\downarrow$

$\underline{a}/\underline{b}$ -> W, t

(3.9) says that given an $\underline{a}$ consisting of anything at all (i.e. W) and a $\underline{b}$ that is [-CASE], there exists in the grammar a rule that allows an $\underline{a}$ that has $\underline{b}$ as its coefficient for SLASH to dominate a W and a trace. The "$\underline{a}/\underline{b}$" notation used here is simply shorthand for the actual feature specification of the mother node which would be [CAT' $\underline{a}$ [FOOT [SLASH $\underline{b}$]]].

Following the analysis in GKPS, the rules responsible for introducing unbounded dependencies are contained in the set of ID rules. Two such rules are given in (3.10):

3.10) a. S -> $\underline{a}$, S/$\underline{a}$

b. S -> PP, VP[there]/PP

By itself, the rule in (3.10a) is responsible for topicalization constructions such as:

3.11) Teddy, we believe will succeed.

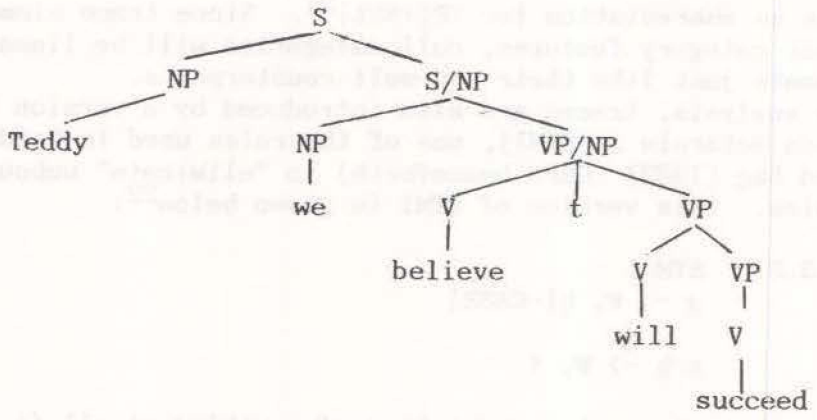
In conjunction with other ID rules and The FOOT feature WH it also accounts for most of the effects of Wh-Movement in a transformational analysis. The rule in (3.10b) is responsible for sentences like (3.12). The feature [there] indicates that the VP is the kind that could take an existential subject as in (3.13):

3.12) In the garden is a fountain.

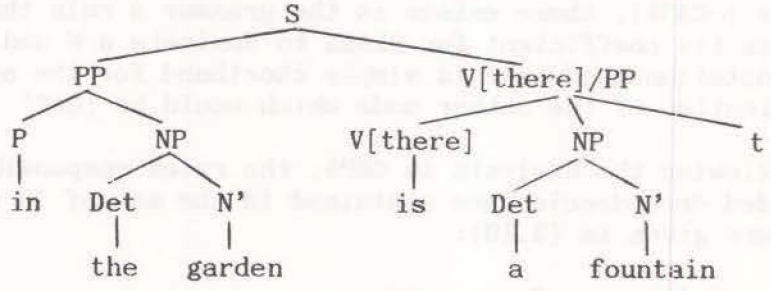
3.13) There is a fountain in the garden.

Since the FOOT feature SLASH takes as its value the category of the "missing" element in an unbounded dependency, this information will be carried through the tree from the point of introduction to the point of elimination by the FFP. Thus given STM1, the rules in (3.10) and the FFP, sentences like (3.11) and (3.12) will be assigned the following structures:

3.14)




3.15)




Since, however, GKPS restrict metarules so as to operate only on ID rules that introduce lexical items, STM1 could not apply in the production of a sentence like:

3.16) John, we believe worked for Kim.

because the rule STM1 would have to apply to would be the rule expanding the S complement of <u>believe</u> as an NP and a VP. This application is blocked since neither NP nor VP is a lexical category. To account for sentences like this, among others, GKPS propose a second STM rule. This rule replaces and generally supersedes the one given in Gazdar (1981) which allowed sentential categories that were missing an NP to be replaced by a VP. Like this rule, STM2 does not involve the introduction of a trace element, rather it allows the remnants of an embedded clause to be "liberated" into a higher clause. This second slash termination metarule is given in (3.17):

3.17) <u>STM2</u>

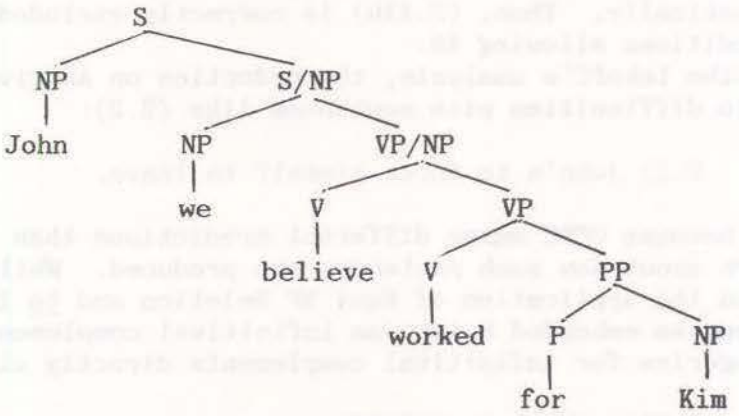
<u>a</u> -> W, <u>b</u>

⇓

<u>a</u>/I -> W, J  where "<u>b</u> -> I, J" is a nonlexical rule

STM2 says that if the grammar has an ID rule which introduced a <u>b</u> and <u>b</u> can dominate I and J, where neither I nor J is a lexical category, then the grammar also has a rule in which <u>b</u> is replaced by J and I is assigned as the value of the mother node's SLASH feature. Given this rule, the sentence in (3.16) would be assigned the structure:

3.18)




Given this account of the distribution of traces in GPSG, statements of the conditions governing the application of AR and CC fall out directly. In the next two sections, I will show how an analysis of the data discussed in chapter two can be devised using a GPSG syntax as a base. Special emphasis will be placed on accounting for dialect variation in a simple and natural way without any ad hoc devices.

3.4 AR in GPSG.

Given a GPSG syntactic framework, the condition on AR in the most liberal dialect is quite easy to state: auxiliaries can contract if they are immediately followed by phonologically non-null material from their own constituents. More restrictive dialects, such as Kaisse's, require an additional condition on possible host elements as well. These dialects will be discussed further below. Notice, however, that I do not attempt to give here a formal statement of what the AR rule looks like. This is because, as we saw in chapter one, both AR and CC are not and can not be syntactic rules themselves. We will return to this issue and how it is predicted by a GPSG framework in section 3.6. Notice also that since cliticization rules are not located in the syntactic component of the grammar they need not be subject to the same types of restrictions as syntactic rules. Just what the general restrictions on cliticization rules are is a topic for future research.

Given this constraint, sentences like (3.19a) will be prevented from undergoing reduction regardless of whether the verb is analyzed as an auxiliary:

3.19) a. I think therefore I am.
b. *I think therefore I'm.

The sentence in (3.19b) is ungrammatical because nothing, not even a trace, follows the auxiliary in its constituent. Even if the verb in (3.19a) is not an auxiliary (which it no doubt isn't), this wording is required on independent grounds to account for sentences like (2.11):

2.11) a. He is, I should think, a bit tired.
b. *He's, I should think, a bit tired.

In this case the verb is an auxiliary by our definition since it takes an XP[PRD] complement (namely an NP) but it still doesn't allow reduction. This is because the material that immediately follows it is not contained

in its clause, rather it is a separate S which is inserted parenthetically. Thus, (2.11b) is correctly excluded by our statement of the conditions allowing AR.

Unlike Lakoff's analysis, the reduction on AR given above does not run into difficulties with sentences like (2.2):

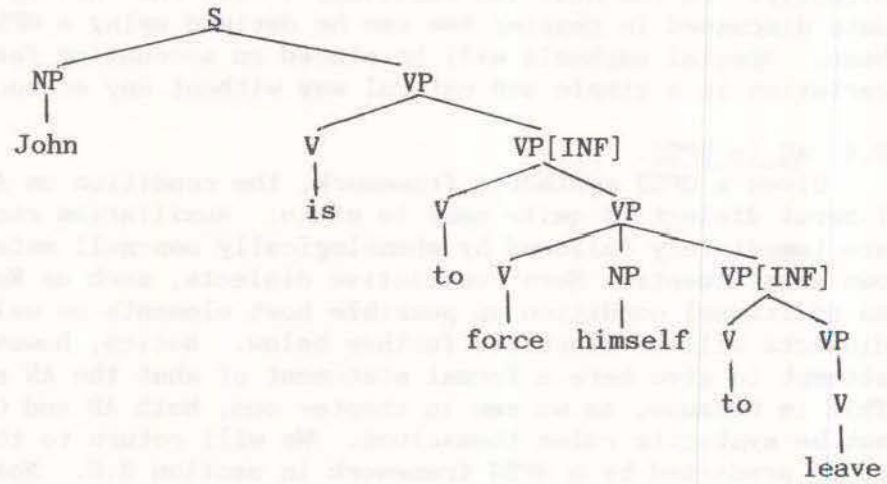
2.2) John's to force himself to leave.

simply because GPSG makes different predictions than transformational grammars about how such sentences are produced. While Lakoff's framework entailed the application of Equi NP Deletion and to Insertion to transform an embedded S into an infinitival complement, verbs in GPSG subcategorize for infinitival complements directly via rules like (3.20):

3.20) VP -> V VP[INF]

Thus the sentence in (2.2) would be assigned the pre-reduction structure in (3.21):

3.21)




Since phonologically non-null material immediately follows the be in its constituent cliticization is possible.

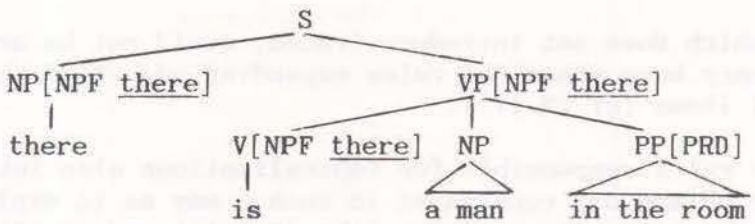
Similarly, under the analysis given in Sag and Klein (1982), so-called There Insertion sentences involve the interaction of subcategorization and agreement rather than string manipulation. Their analysis relies on rules like the following, adapted from the original X-bar notation (Sag and Klein, p. 103):

3.22) < 7; NP[NPF a] -> a >, where a $\in$ {it, there}
3.23) < 12; VP -> V[-PRP, NPF there], NP, XP[PRD] >, where V[12] = {be}

NPF is a type of agreement feature which insures that dummy NP's appear in the appropriate structures. Since agreement is stated between subject NP's and their VP's[23] and subsequently carried through the tree by the HFC, these rules result in pre-reduction structures such as (3.24) for sentences like (2.1ai):

2.1) a. i. There's a man in the room.

3.24)

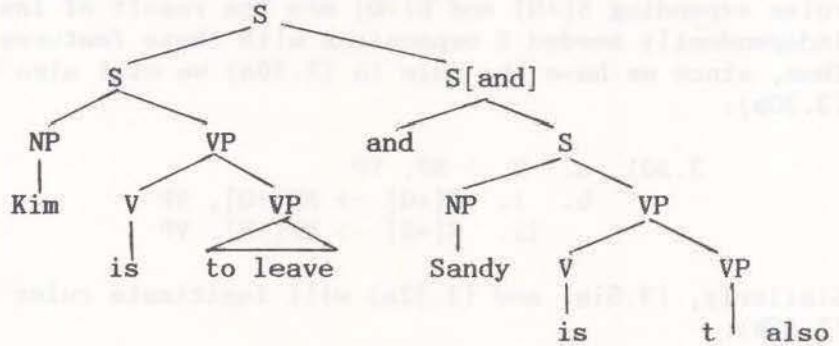



Unlike some transformational analyses, there is no disturbance in the material following the auxiliary and thus no need, as with Kaisse (1983a), to refer to more than one level of structure in order to account for the ability of AR to apply here.

While Equi and existential there sentences do not involve phonologically null elements, the GPSG equivalent of VP Deletion, as we have seen, does. Therefore, if a VP that is assigned the feature [+NUL] immediately follows an auxiliary that auxiliary cannot undergo AR. Thus a sentence like (2.1di) will not be grammatical since it is assigned the pre-reduction structure given in (3.8):

2.1) d. i. *Kim is to leave and Sandy's, also.

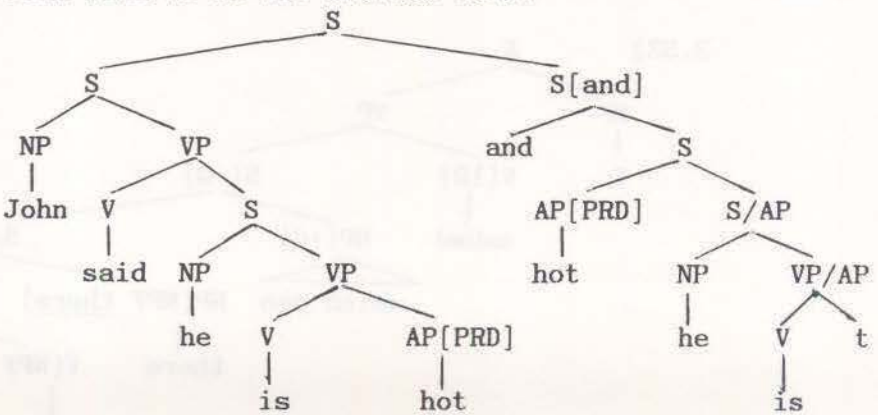
3.8)




AR can also be blocked from applying in topicalized sentences since STM1 will be used to eliminate the unbounded dependency. Thus a sentence like (3.25a) could not undergo reduction in the second conjunct since it would be assigned the structure in (3.26), with a phonologically null element (a trace) after the auxiliary:

3.25) a. John said he is hot and hot he is!
b. *John said he is hot and hot he's!

3.26)



STM2, which does not introduce traces, could not be used to eliminate the dependency here since the rules expanding adjective phrases introduce lexical items (cf (3.17)).

The rules responsible for topicalizations also interact with the FOOT feature WH and our constraint in such a way as to explain the ungrammaticality of sentences like (2.laii) and (3.27):

2.1) a. ii. *I asked which men there're in the room.
3.27)*The restaurant in whose cellar that wine's will be the most popular.

In GKPS, the ID rules which introduce embedded questions and relative clauses are the following:

3.28) VP -> V[18], S[+Q]
3.29) NP -> NP, S[+R]

where Q is the value WHMOR takes for interrogatives and R the value it takes for relatives. [+Q] and [+R] are used as abbreviations for the features [WH AGR [WHMOR Q]] and [WH AGR [WHMOR R]] respectively. The rules expanding S[+R] and S[+Q] are the result of instantiating independently needed S expansions with these features by the FFP[24]. Thus, since we have the rule in (3.30a) we will also have the rules in (3.30b):

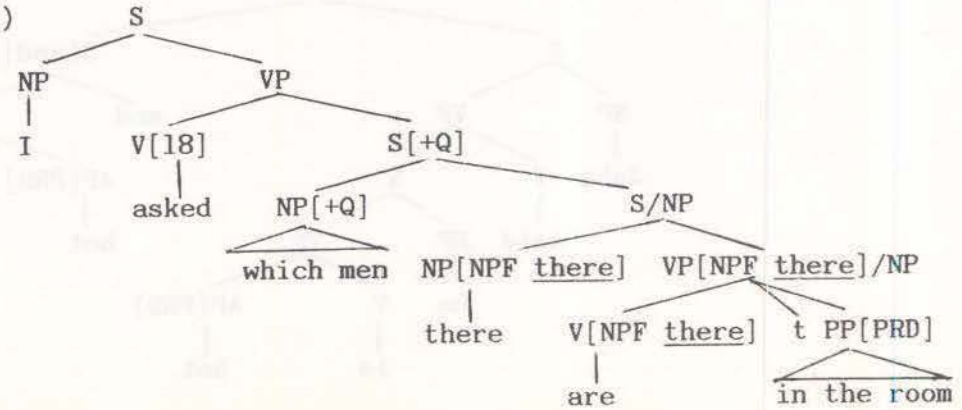
3.30) a. S -> NP, VP
b. i. S[+Q] -> NP[+Q], VP
ii. S[+R] -> NP[+R], VP

Similarly, (3.31a) and (3.32a) will legitimate rules like (3.31b) and (3.32b):

3.31) a. S -> PP, VP[there]/PP
b. i. S[+Q] -> PP[+Q], VP[there]/PP
ii. S[+R] -> PP[+R], VP[there]/PP
3.32) a. S -> a, S/a
b. i. S[+Q] -> a[+Q], S/a
ii. S[+R] -> a[+R], S/a

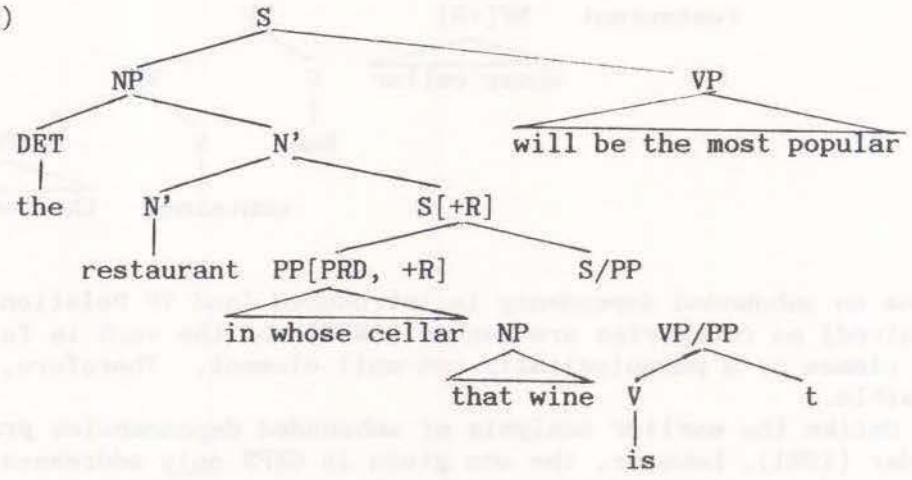
Given the rule in (3.32bi), the pre-reduction structure for (2.laii) would be as follows, where a = NP:

3.33)



Since the unbounded dependency introduced in connection with the NP which men is eliminated by STM1 a trace is left. Since this trace has all the category features of an NP, it will be linearized into the position following the verb are by the LP statements, thus preventing AR from applying. Sentences such as (3.27) are blocked for similar reasons, as we can see from the structure in (3.34):

3.34)

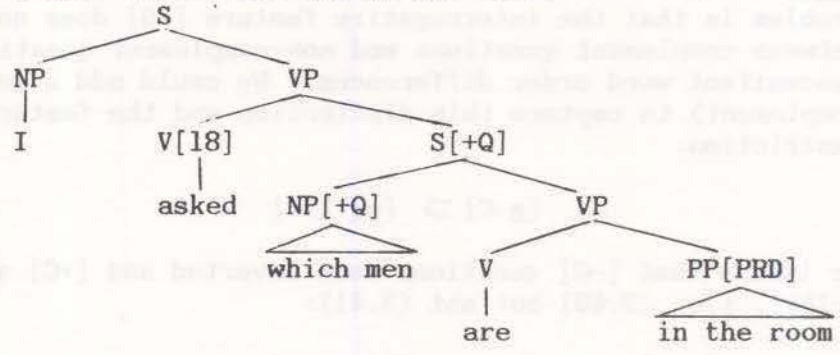



The relative clause expansion used here is the one given in (3.32bii) with a = PP.

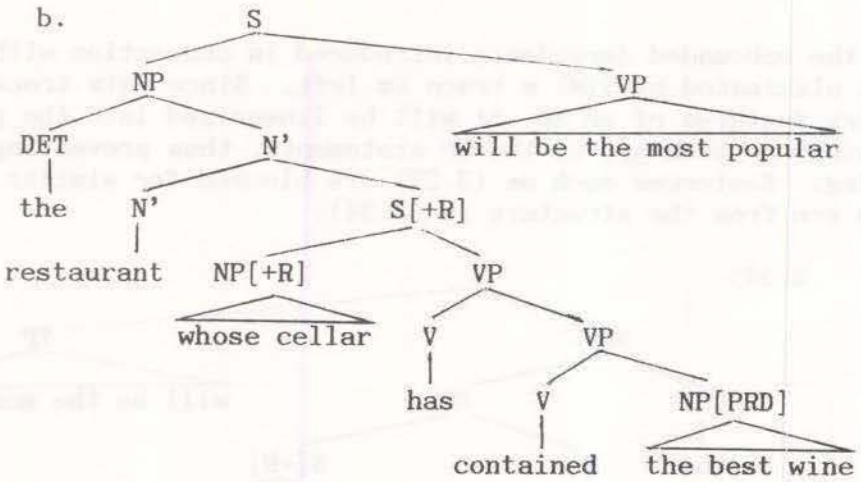
Notice, however, that superficially similar sentences such as (3.35a) and (3.36a) will allow AR to apply since their embedded sentences are expanded by the rules in (3.30b), as shown in (3.35b) and (3.36b):

3.35) a. I asked which men are in the room.

b.




3.36) a. The restaurant whose cellar has contained the best wine will be the most popular.




Since no unbounded dependency is introduced (and VP Deletion is not involved) no categories are marked [+NUL] and the verb is followed in its own clause by a phonologically non-null element. Therefore, AR is possible.

Unlike the earlier analysis of unbounded dependencies presented in Gazdar (1981), however, the one given in GKPS only addresses the issues of embedded sentences and relative clauses; no mention is made of root wh questions such as (3.37)-(3.39):

3.37) Who is Pita?
3.38) In which garden is a statue of Pita?
3.39) In which park did Pita say the concert is?

In fact, given the assumptions in GKPS it is difficult to see how they could account for such sentences in a reasonable and regular way. One problem is that the interrogative feature [+Q] does not distinguish between complement questions and non-complement questions and their concomitant word order differences. We could add a feature [± C] (i.e complement) to capture this distinction and the feature co-occurrence restriction:

$$[\underline{a}\ C] \supset [-\underline{a}\ INV]$$

to insure that [-C] questions were inverted and [+C] question were [-INV], i.e. (3.40) but not (3.41):

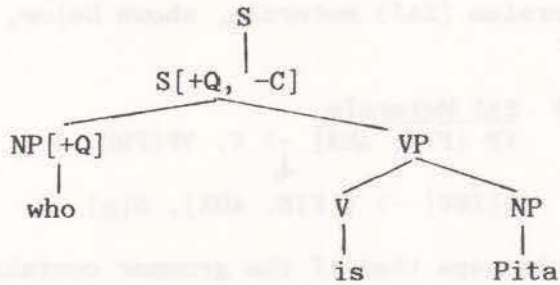
3.40) a. Who can Pita see?
b. I know who Pita will see.
3.41) a. *Who Pita can see?
b. *I know who will Pita see.

but even this would not give us a completely adequate account of root wh questions. This is because the FOOT feature [+Q] works the same way as the FOOT feature SLASH (GKPS, p. 54) and, as such, must be specifically introduced by an ID rule. Thus, if we wish to take advantage of the prediction of feature instantiation, as we did with relative clauses and embedded questions, we would have to propose a rule like the following:

3.42) S -> S[+Q, -C]

Notice, however, that this rule produces some very peculiar tree structures:

3.43)




There is no independent justification for the extra S node dominating the S[+Q, INV], rather its only reason for existence is to allow us to introduce the [+Q] feature.

We could, of course, simply list each of the rules expanding an S[+Q, -C] separately in the grammar as shown below:

3.44) a. S[+Q, -C] -> NP[+Q], VP
b. S[+Q, -C] -> PP[+Q], VP[there]/PP
c. S[+Q, -C] -> a, S[INV]/a
etc.

but this would result in a great deal of redundancy in the ID rules and fails to capture generalizations about the feature system and the structure of root wh sentences. It would be better if the grammar somehow predicted the existence of the rules in (3.44).

A possible solution to these problems would be to give up GKPS's stipulation that metarules may only apply to lexical ID rules and introduce root wh question expansions via the following[25]:

3.45) S[+Q] -> H, W
↓
S -> H[INV], W

where the default value for WH is assumed to be null. This metarule says that if the grammar has a rule that expands an S[+Q] as a head and its complements, then the grammar will also have a rule that expands a regular S in the same way except that its head will be marked with the inversion feature[26]. Thus, since the grammar will have the rules expanding S[+Q] given in (3.30bi), (3.31bi) and (3.32bi) the grammar will also have the rules in (3.46)[27]:

3.46) a. S -> NP[+Q], VP[INV]
b. S -> PP[+Q], VP[there, INV]/PP
c. S -> a, S[INV]/a

Recall that the notation "a/b" is simply an abbreviation for [CAT' a [FOOT [SLASH b]]] and that head daughters are chosen on the basis of bar level and syntactic category. Since having a value for SLASH in no way affects the category or bar level features of a node, an a/b can qualify

as the head of a phrase as long as the other criteria are met. Thus the VP in (3.46b) and the S in (3.46c) will be assigned the inversion feature by (3.45) despite the fact that they are slashed categories (i.e. a slashed VP is still a VP, etc.).

The rule in (3.46c) will interact with the output of the Subject Auxiliary Inversion (SAI) metarule, shown below, to produce sentences like (3.39)[28]:

3.47) SAI Metarule

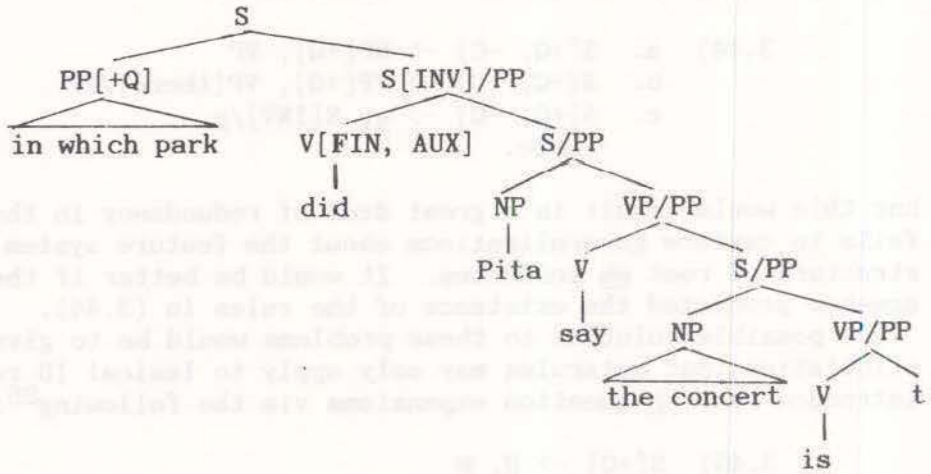
VP [FIN, AUX] -> V, VP[PRD, a]

↓

S[INV] -> V[FIN, AUX], S[a]

The SAI metarule says that if the grammar contains an ID rule expanding a finite verb phrase as an auxiliary verb and its VP complement, then the grammar also has a rule expanding an S with the inversion feature as a finite auxiliary verb and an S with the same subcategorization features as the VP complement of the auxiliary in the input rule. Thus, the structure for a sentence like (3.39) would be as in (3.48), where a in (3.46c) takes the value PP:

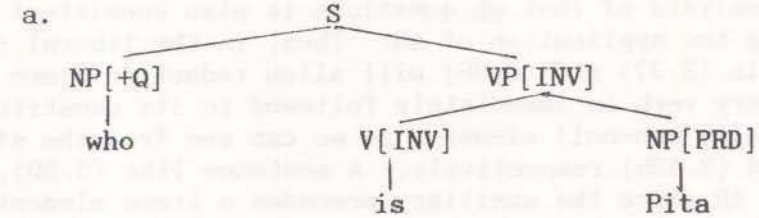
3.48)



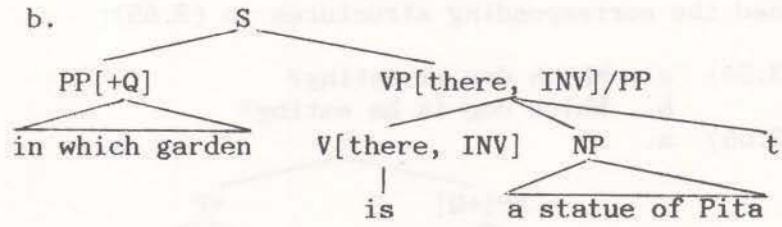

Since we are assuming that unless otherwise specified no value for the interrogative feature is assigned, we do not need to worry about introducing the nodes which expand as root wh questions in other ID rules, as we did with embedded questions. Also, if we assume that the default value for the feature [INV] is negative, we can insure that embedded questions like (3.41b) can not be generated (alternately we could specify the value [-INV] in the rule that introduces embedded questions itself, i.e. (3.28)).

As shown above, the rule in (3.45) also assigns the [INV] feature to the VP's in (3.46a,b). By the HFC this feature will be passed on to the V's these VP's dominate to produce structures like (3.49):

3.49) a.




b.




This would be fine except for the feature co-occurrence restriction proposed in Gazdar, Pullum and Sag (1981) shown below:

3.50) [INV] ⊃ [AUX, FIN]

(3.50) says that if something has the feature [INV], then it will also be an auxiliary and be finite. Given this restriction we would not be able to produce a sentence like (3.51):

3.51) Who loves Pita?

since loves is not an auxiliary. This FCR was proposed to prevent sentences like (3.52) from being legitimized by the SAI metarule:

3.52) *Loves Pita to sing?

If, however, we formulate the SAI metarule as in (3.47), with the features [AUX] and [FIN] specifically stated on the verb, sentences like (3.52) will be blocked and the FCR in (3.50) made superfluous. Thus we can dispense with (3.50) without making false predictions[29].

Nor will allowing the feature [INV] to sometimes appear on V's that don't begin a sentence interfere with the treatment of morphological irregularities such as the following:

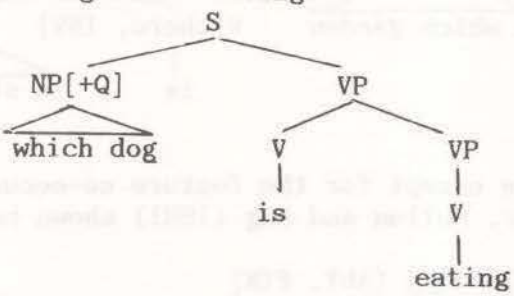
3.53) a. *I amn't going.
b. *Amn't I going?
c. *I aren't going.
d. Aren't I going?

As Gazdar, Pullum and Sag observe (p. 31), this paradigm can be accounted for simply by stipulating in the lexicon that the first person singular present tense copula has no [-INV, +n't] form and that its [+INV, +n't] form is aren't. Since the only time that V's that don't begin a sentence are marked as being inverted is in connection with wh words or phrases, and since wh words and phrases are always third person never first person, this lexical restriction will not be affected and we will not incorrectly predict that sentences like (3.53c) are grammatical.

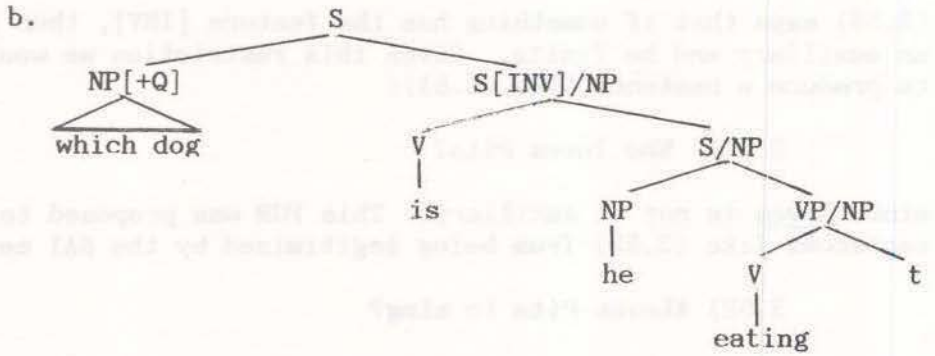
This analysis of root wh questions is also consistent with the facts surrounding the application of AR. Thus, in the liberal dialect, the sentences in (3.37) and (3.38) will allow reduction since in each case the auxiliary verb is immediately followed in its constituent by a phonologically non-null element, as we can see from the structures in (3.49a) and (3.49b) respectively. A sentence like (3.39), however, will not permit AR since the auxiliary precedes a trace element, as shown in (3.48). Similarly, both the sentences in (3.54) will allow AR since they are assigned the corresponding structures in (3.55):

3.54) a. Which dog is eating?
b. Which dog is he eating?

3.55) a.




b.




The importance of the difference between these two structures will become apparent in the following section.

3.4.1 Conservative Dialects.

While the constraint on AR given at the beginning of section 3.3 correctly predicts the facts about the dialect described there, other dialects are not quite as "liberal" with regard to where they permit reduction to occur. The dialect described in Kaisse (1981), for example, differs from more liberal dialects in that it does not allow reduction in sentences such as the following, repeated here from section two:

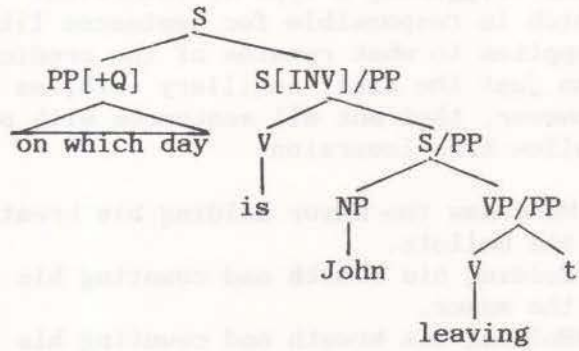
3.56) Not only's Louis smart, he's also a varsity rower.
3.57) On which day's John leaving?
3.58) a. Speaking tonight's a famous reporter.
b. Speaking tonight's been a famous reporter.
3.59) Under this slab's buried Joan of Arc.

Since such sentences are perfectly acceptable in the more liberal dialect, their ungrammaticality here cannot be the result of phonologically null elements being positioned after the auxiliary. There

must, therefore, be some other type of constraint at work in these cases. In order to determine what this constraint would be, let us briefly consider the structures such sentences could be assigned in GPSG and what they have in common.

Sentences similar to (3.57) have already been discussed in connection with root wh questions; given the rule in (3.46c), (3.57) would be assigned a structure very much like the one in (3.55b) above:

3.60)



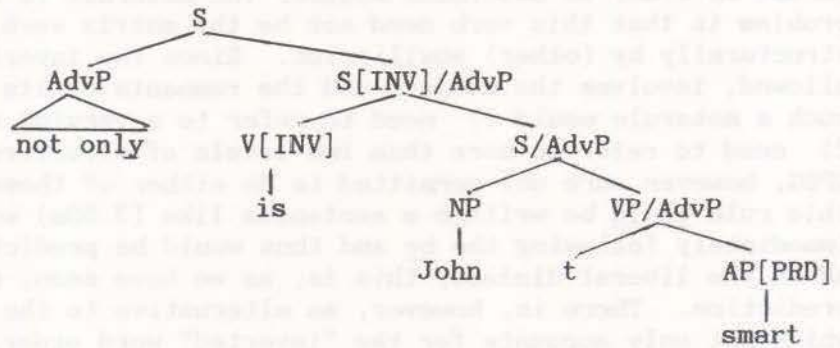

The rest of the sentences listed above, however, have not (to my knowledge) been previously addressed within the GPSG framework.

The analysis of sentences like (3.56) in GPSG, as we shall see, is fairly straightforward. Kaisse, following Emonds (1976), analyses such sentences as being derived from the deep structure in (3.61) by Negative Preposing and SAI:

3.61) Louis is not only smart, he's also a varsity rower.

Assuming the basic correctness of this choice of positioning for the adverb phrase, the corresponding GPSG structure would be as follows:

3.62)




where the "preposing" is achieved by the familiar rules for topicalization. The inverted word order can be guaranteed by means of a FCR on slash introductions such as the following[30]:

$$[\text{SLASH } \underline{a}[+\text{Neg}]\ ] \supset [\text{INV}]$$

Notice that the auxiliary in this sentence does indeed meet the conditions on AR outlined in the preceding section; thus, its ability to reduce in the liberal dialect is explained.

The analyses of sentences like (3.58) and (3.59), however, are a bit more involved. On a transformational analysis (3.58a) would be assigned the deep structure given in (3.63):

3.63) A famous reporter is speaking tonight.

(3.58a) would then be derived by applying a preposing rule to the participal phrase thus triggering a type of inversion. This inversion differs from SAI, which is responsible for sentences like (3.56) and (3.57), in that it applies to what remains of the predicate after preposing rather than just the first auxiliary verb, as we can see from (3.58b). Notice, however, that not all sentences with preposed participal phrases allow this inversion:

3.64) a. Mary saw the mayor holding his breath and counting his ballots.
b. Holding his breath and counting his ballots Mary saw the mayor.
c. *Holding his breath and counting his ballots saw Mary the mayor.

Nor are participles the only preposed phrases which do trigger it:

3.65) a. Snowflakes of all shapes and sizes had fallen from the sky.
b. From the sky had fallen snowflakes of all shapes and sizes.

Since, in transformational terms, preposing the complement of some verbs can result in this type of inversion while preposing the complement of other verbs does not, a metarule to produce such structures in GPSG would need to refer to the verb which subcategorizes for the topicalized phrase in order to determine whether the metarule is applicable. The problem is that this verb need not be the matrix verb, it can be preceded structurally by (other) auxiliaries. Since the inverted order, when allowed, involves the subject and the remnants of its verbal complement, such a metarule would 1) need to refer to a varying number of levels and 2) need to refer to more than two levels of structure. Metarules in GPSG, however, are not permitted to do either of these. Furthermore, if this rule could be written a sentences like (3.58a) would have a trace immediately following the be and thus would be predicted not to undergo AR in the liberal dialect; this is, as we have seen, not a correct prediction. There is, however, an alternative to the metarule approach which not only accounts for the "inverted" word order, it also makes the correct predictions about the positioning of traces.

First of all, notice that verbs like be and fall (but not saw) have something in common other than the fact that when their complements are topicalized this inverted order is found: both be and fall allow existential subjects. Thus, alongside sentences like (3.56) and (3.65) we find sentences like:

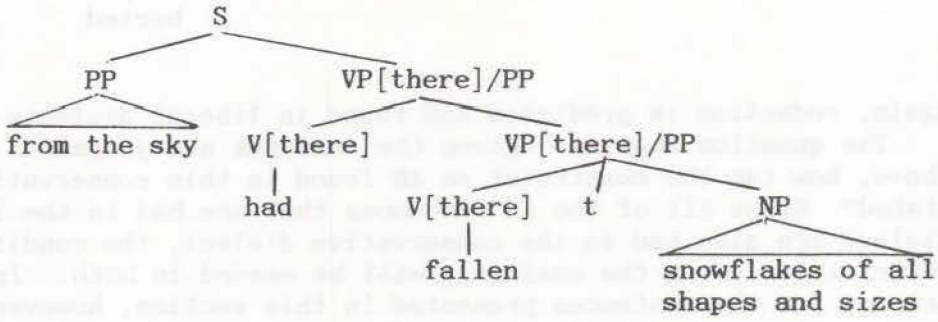
3.66) There is a famous reporter speaking tonight.
3.67) There had fallen from the sky snowflakes of all shapes and sizes.

Since the matrix VP's in these sentences carry the feature [NPF there] in agreement with their subjects, (3.67) can have unbounded dependencies introduced by the ID rule in (3.10):

3.10) S -> PP, P[there]/PP

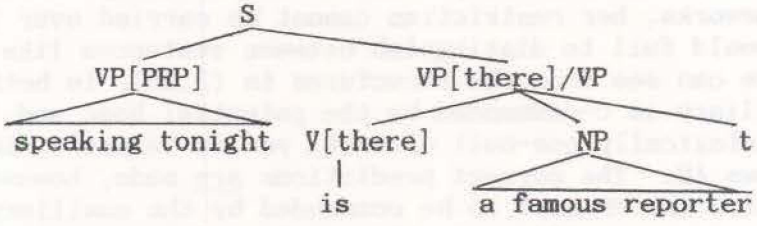
Thus, the sentence in (3.65b) could be assigned the structure:

3.68)

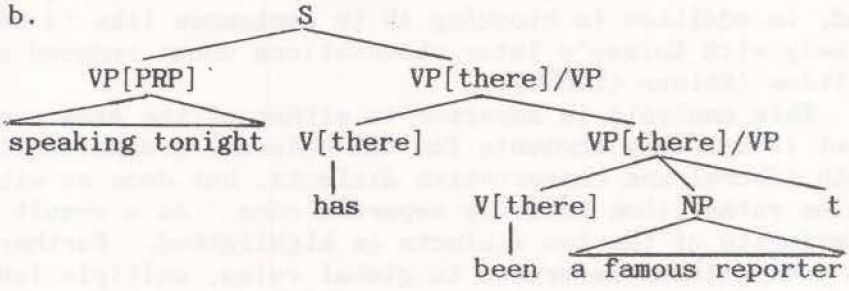



Sentences like (3.58) can be accounted for by generalizing the rule in (3.10) so as to apply to VP[PRD] as well, yielding structures such as:

3.69) a.




b.



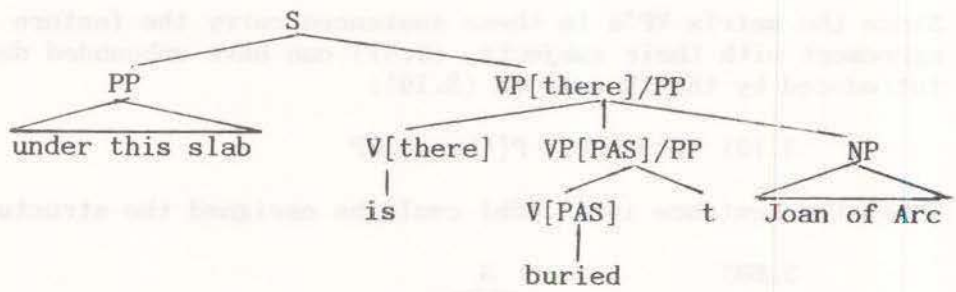

Notice that in each case the initial auxiliary verb is followed in its constituent by a phonologically non-null element. Thus this analysis, unlike the transformational one outlined above, will allow AR to apply here in liberal dialects.

A similar analysis can be proposed for sentences like (3.59). Since there will be ID rules in the grammar licensing the occurrence of sentences like:

3.70) There is buried under this slab Joan of Arc.

then, given the rule in (3.10b), sentences like (3.59) will also be admitted. A likely structure for such a sentence is given in (3.71)[31]:

3.71)




Again, reduction is predicted and found in liberal dialects.

The question then is: given the analyses and judgments presented above, how can the constraint on AR found in this conservative dialect be stated? Since all of the AR sentences that are bad in the liberal dialect are also bad in the conservative dialect, the condition on material following the auxiliary will be needed in both. In order to account for the sentences presented in this section, however, the conservative dialect will also need a means of restricting the set of elements that can serve as host to the cliticization. The constraint proposed by Kaisse (1983b), recall, restricted the host to NP's which c-command the auxiliary. Due to the different structures assigned by our frameworks, her restriction cannot be carried over into this work since it would fail to distinguish between sentences like (3.54a) and (3.54b). As we can see from the structures in (3.55), in both sentences the auxiliary is c-commanded by the potential host and followed by a phonologically non-null element, yet in Kaisse's dialect only (3.54a) allows AR. The correct predictions are made, however, if we instead require the NP host to be commanded by the auxiliary. This in effect prevents an auxiliary in this dialect from attaching outside of its own S and, in addition to blocking AR in sentences like (3.54b), ties in very nicely with Kaisse's later observations about reduced auxiliaries and 2P clitics (Kaisse (1983b)).

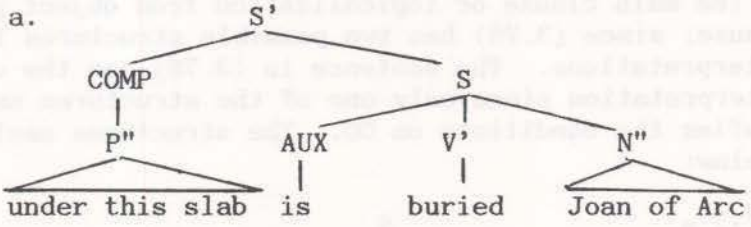
This analysis is superior to either of the ones proposed by Kaisse in that it not only accounts for the relevant grammaticality judgments in both liberal and conservative dialects, but does so with a related set of rules rather than entirely separate ones. As a result the underlying similarity of the two dialects is highlighted. Furthermore, all of this is done without reference to global rules, multiple levels of structure or transformations—extremely powerful devices which are simply not needed in a GPSG syntax. Finally, this approach has the added benefit of giving insights into judgments in the conservative dialect that Kaisse herself could not explain.

Kaisse (1983a) notes (p. 109) that some speakers do not find inversion sentences with preposed PP's such as (3.59) as bad as inversion sentences with preposed participles (cf (3.58)), a fact that she attributes to the relative NP-ness of the two types of phrases. If, however, the greater degree of acceptability found with sentences like (3.59) is a result of the NP-ness of the PP host, we would also expect sentences like (3.57) to be relatively more acceptable as well, since they also have PP hosts. Such is apparently not the case in these dialects. Nor can Kaisse rely on the difference in following context to differentiate the two sentence types since sentences like (3.72) receive similar judgments (Kaisse (1983a), p. 109):

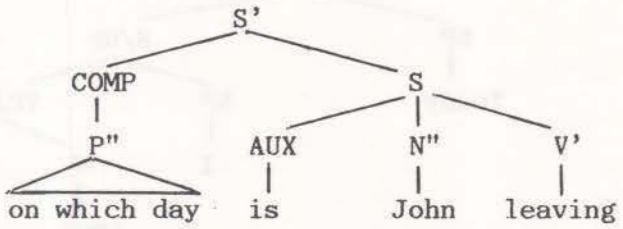
3.72) ?Implicit in your statement's the idea that men are inferior.

A GPSG syntax, however, provides a straighforward explanation of this difference in host status: the PP in (3.57), as we can see from (3.60), fails the condition on hosts on two counts 1) it is not an NP and 2) it is not commanded by the auxiliary seeking to reduce. The PP in (3.59), on the other hand, is commanded by the reducing auxiliary (cf (3.71)). Thus if the NP host condition is relaxed for the speakers in question so as to include PP's, there would be nothing to prevent the auxiliary in (3.59) (or (3.72)) from reducing while AR in (3.57) would still be blocked on structural grounds. No such structural distinction between PP hosts is possible in Kaisse's analysis, however, since in her theory all preposed material is inserted into COMP:

3.73) a.




b.

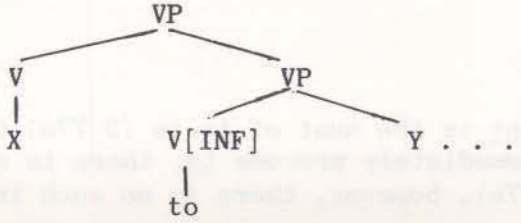



Since the structures in (3.73) are completely analogous (except for the different status of the material following AUX which, as we've seen, cannot be the cause of the grammaticality distinction) no competing explanation of these facts is available.

## 3.5 CC in GPSG.

As was the case with AR, the syntactic conditions governing the application of CC are relatively easy to state assuming we use a GPSG syntax. Within the majority dialect this condition is as follows: contraction is possible only if the node introducing the trigger verb[32] is the aunt of the node introducing to and they are linearly adjacent. Put more simply, contraction is only possible in structures such as the following[33]:

3.74)



Such a definition entails that in order for CC to take place the triggering verb must c-command the to. Thus, sentences such as (2.6c) and (2.7b), repeated below, are ungrammatical since in these cases want does not c-command to:

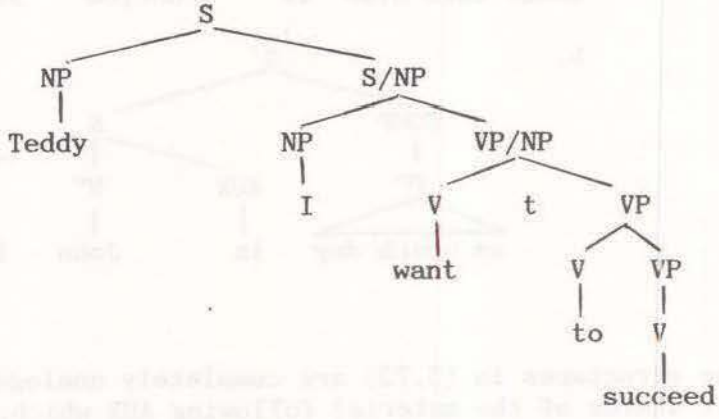
2.6) *It seems like to wanna regret what one does not have.
2.7) *I don't want anyone who continues to wanna stop wanting.

This condition also provides an account of the ambiguity constrasts found between pairs like the following:

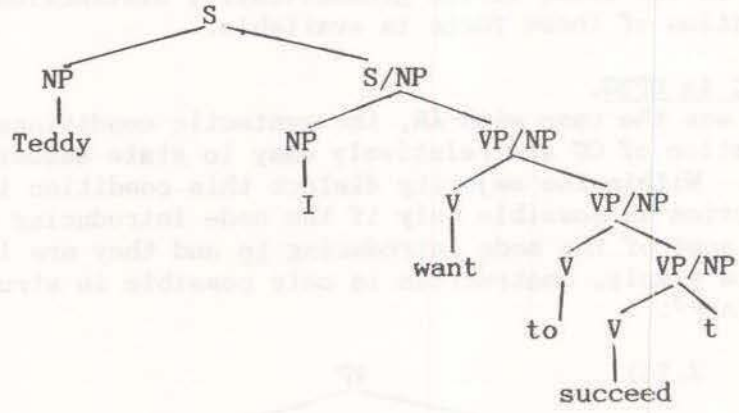
3.75) Teddy, I want to succeed.
3.76) Teddy, I wanna succeed.

The sentence in (3.75) could involve either topicalization from object position in the main clause or topicalization from object position in the embedded clause; since (3.75) has two possible structures it also has two possible interpretations. The sentence in (3.76), on the other hand, has only one interpretation since only one of the structures assigned to (3.75) satisfies the conditions on CC. The structures assigned to (3.75) are shown below:

3.77) a.




b.

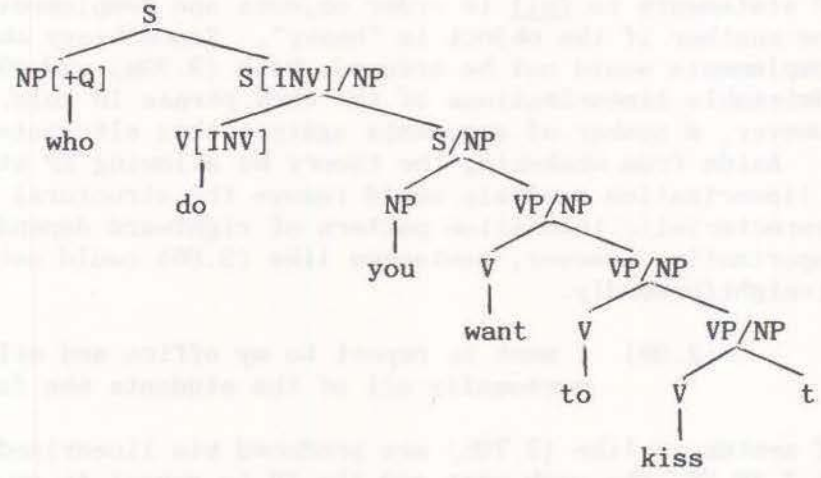



Though the trigger want is the aunt of to in (3.77a) CC is not possible since want does not immediately precede to, there is a trace intervening. In (3.77b), however, there is no such intervening trace and

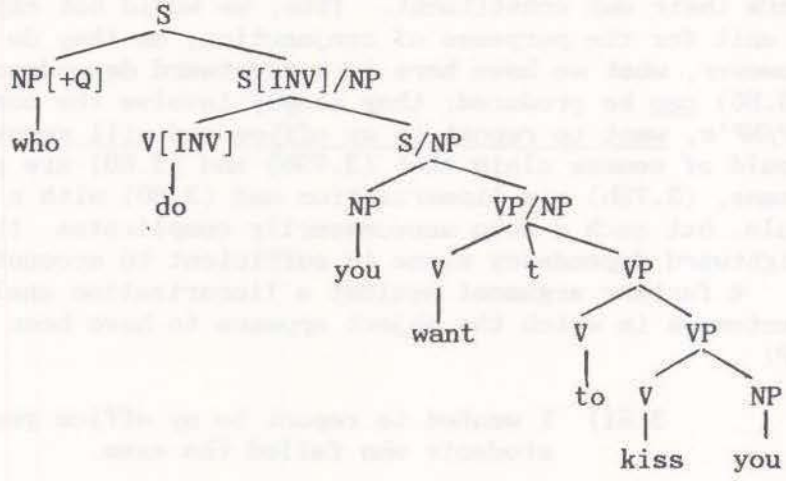
CC is indeed possible. Thus the sentence in (3.75) has the same interpretation as the version of the unreduced sentence in which the object in the embedded clause is what is topicalized, i.e. the I want to succeed Teddy reading.

Given the analysis of root wh questions discussed in the preceding section, a similar treatment is available for sentences like (2.24b) and (2.25b). The structures corresponding to these sentences are shown in (3.78):

3.78) a.




b.




Since the verb want in (3.78a) immediately precedes to and is also its aunt CC is possible and, thus (2.24b) is grammatical. CC is blocked for (2.25b), however, because of the intervening trace shown in (3.78b). Notice that unlike recent transformational accounts, there is no need to distinguish here between different types of traces (i.e. case marked versus non-case marked) since the rules responsible for wh sentences do not reapply for each S (i.e. are not successive cyclic) and thus do not overgenerate null elements.

All of the cases of unbounded dependency discussed thus far have involved structures in which an element which shares the category value of the SLASH feature is linearized to the left of the category that bears this feature. English, however, also allows sentences with rightward

dependencies, an example of such a sentence being the one in (3.79b) in which an object NP is shifted to the end of its verb phrase:

3.79) a. I want all of the students who failed the exam to report to my office.
b. I want to report to my office all of the students who failed the exam.

Another possible treatment of sentences like (3.79) would be to allow the LP statements to fail to order objects and complements with respect to one another if the object is "heavy". Since heavy objects and complements would not be ordered, both (3.79a) and (3.79b) would be admissable linearizations of the verb phrase ID rule. There are, however, a number of arguments against this alternate analysis.

Aside from weakening the theory by allowing LP statements more power, a linearization analysis would remove the structural basis for the characteristic intonation pattern of rightward dependencies. More importantly, however, sentences like (3.80) could not be accounted for straightforwardly:

3.80) I want to report to my office and will speak to personally all of the students who failed the exam.

If sentences like (3.79b) are produced via linearized ID rules such as VP -> V VP NP, the verb want and the VP to report to my office would not form their own constituent. Thus, we would not expect them to operate as a unit for the purposes of conjunction, as they do in (3.80). If, however, what we have here is a rightward dependency, sentences like (3.80) can be produced; they simply involve the conjunction of two VP/NP's, want to report to my office and will speak to personally. One could of course claim that (3.79b) and (3.80) are produced by different means, (3.79b) via linearization and (3.80) with a rightward dependency rule, but such a move unnecessarily complicates the grammar since rightward dependency alone is sufficient to account for both.

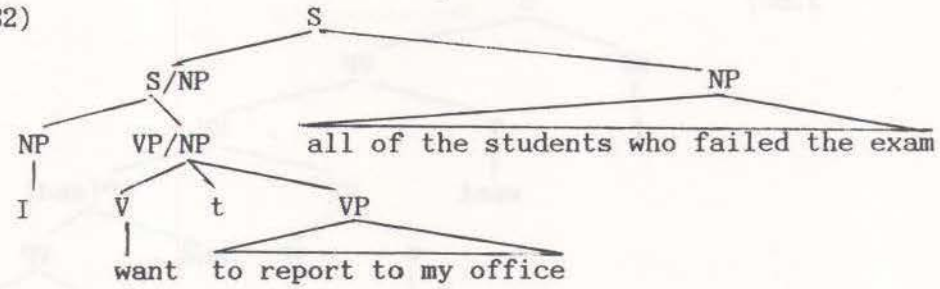
A further argument against a linearization analysis is provided by sentences in which the object appears to have been shifted outside of its VP:

3.81) I wanted to report to my office yesterday all the students who failed the exam.

On the interpretation where it was yesterday that I wanted X to happen, the adverb yesterday will be modifying the entire sentence. We will therefore have an ID rule like S -> S, Adv. Since the NP object in (3.81) is outside of the sentential adverb, it must also be outside of its own VP. Thus a linearization analysis cannot account for the clause order in (3.81) under this interpretation.

Given that sentences such as (3.79b) involve dependencies, the rules we have discussed thus far will provide structures like the following, assuming the correct linearization principle[34]:

3.82)




Since the LP statements will place the trace of the object NP between the want and to of the matrix VP this analysis correctly predicts that sentences like (3.79b) will not allow CC:

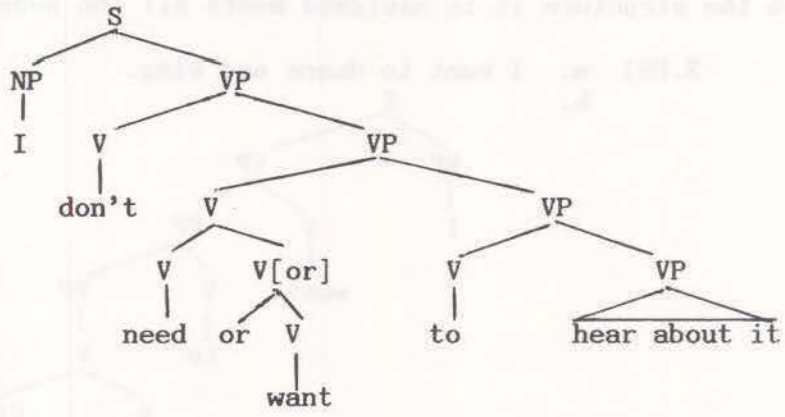
3.83) *I wanna report to my office all of the students who failed the exam.

This analysis of CC also predicts the failure of contraction in sentences with conjoined verbs such as (2.9a), repeated below:

2.9) a. I don't need or want to hear about it.

The structure assigned to such a sentence is shown in (3.84):

3.84)

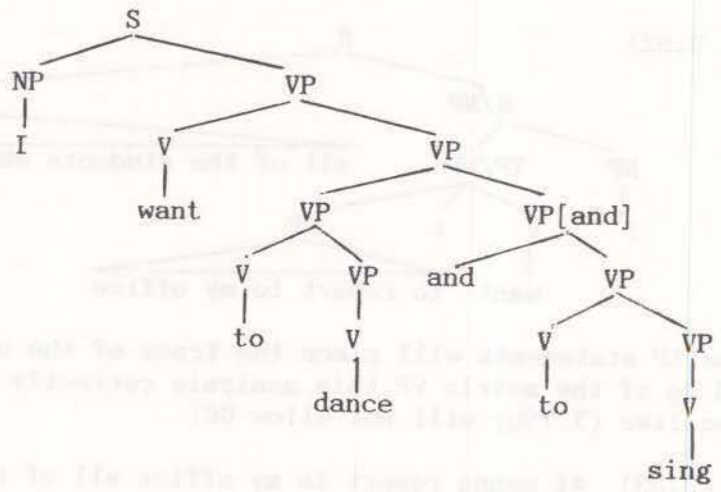



Since the V dominating want does not c-command the to, it is not its aunt and contraction is not possible there. Nor can contraction apply at the next higher level since that V does not introduce a trigger verb[35]. Since we are assuming that CC is not a syntactic rule we need not worry about possible ordering paradoxes between ID rules and/or metarules[36].

A similar treatment is possible for sentences like (2.8a) in which verb phrases are conjoined:

2.8) a. I want to dance and to sing.

3.85)



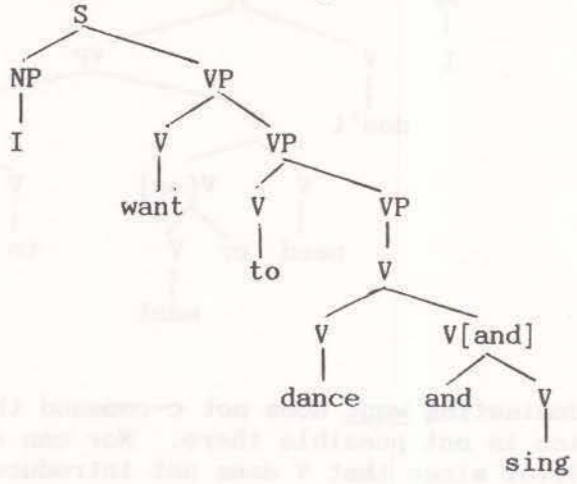

Since the node dominating want is not the aunt of the node dominating the (first) to in this structure, CC is blocked. Thus sentences like (2.8b) are correctly predicted to be ungrammatical:

2.8) b. *I wanna dance and to sing.

Notice, however, that a sentence like (3.86a) will allow reduction since the structure it is assigned meets all the necessary requirements:

3.86) a. I want to dance and sing.

b.




Thus the analysis presented above correctly handles the entire range of contraction and conjunction facts.

3.5.1 Liberal dialects.

Like AR, judgments about CC are subject to variation from dialect to dialect. Unfortunately the scope of this variation is not quite as well documented as with AR, thus making generalizations difficult. Some speakers seem to be more "liberal" than most with regard to CC in that the presence of trace between a to and its aunt does not block

contraction. Such speakers find sentences such as (3.87) perfectly acceptable (though perhaps marked as very informal):

3.87) Who do you wanna kiss you?

Whether traces in such dialects behave differently with respect to AR as well has not, to my knowledge, ever been discussed; thus there is no way of telling whether the transparency in (3.87) is a fact about traces in general in these dialects or simply a fact about CC.

If the latter is the case then we need only change the condition on CC to something like the following: CC is possible in a configuration XYZ iff X is the aunt of Z and Y is phonologically null. Since the function of rules like CC and AR seems to be to make phonological units out of syntactically distinct items, it does not seem unreasonable that in some dialects the syntactic aspect would be the deciding factor in contraction while in others the phonological aspect would. Thus the two dialects differ only in minor details.

If, on the other hand, this transparency is a fact about traces as a whole for these speakers we have two options: we could either modify both AR and CC so as to allow traces in the relevant sites, or we could adopt a position for these dialects only similar to the one taken in GKPS and fail to have STM1 introduce traces. Which view of traces is correct and which modification of the theory is preferred is an empirical question in need of further research.

3.6 Ordering of AR and CC.

In section one I presented arguments against a morphological or phonological treatment of AR and CC as well as several preliminary arguments against analyzing either one as a syntactic rule. These latter arguments revolved around the fact that neither AR nor CC interacts crucially with any other syntactic rule and the fact that rules like AR and CC have very different functions than syntactic rules. The view of grammatical organization that I adopted there was one in which rules like AR and CC were contained in a separate component of the grammar reserved for cliticizations. While this type of highly constrained non-syntactic treatment of cliticization is consistent with most transformational analyses of AR and CC, it is in no way predicted by them. As we shall see, a non-syntactic treatment is not only consistent with the GPSG system, it is actually required by it.

Consider first of all the type of statement a rule like AR would require if it were part of a GPSG syntax. As we saw in section (3.4), AR in the liberal dialect is possible in many different structures onto many different elements -- NP's, AdvP's, VP's etc. Thus, there will be no general way of specifying what the result of cliticization would look like. This plus the fact that rules in GPSG can in general only refer to two levels of structure at a time means that each of the eligible structures, if characterizable at all, would require a separate rule. This in turn suggests the possibility of dialects which contain some of these rules but not others, or which contain some rules with quite different constraints. Thus it would be logically possible for a dialect to have rules which allow AR in some cases only if the auxiliary is not followed by a phonologically non-null item and in other cases only if it is. The fact that the actual rules for AR in the dialects studied thus far all share the same constraint on following context would be treated

as accidental. Since this is obviously not a desirable prediction we need some principled way of excluding this possiblility.

A second and even more damaging problem for a syntactic treatment of AR and CC in GPSG hinges on the fact that PS rules act as node admissibility conditions. Given this interpretation, it makes no sense to talk about derivations or ordered relationships among PS rules. Thus, a phrase is well formed if there is a PS rule for that phrase which allows the node it expands to dominate the categories it does in the proper order. This means that if AR or CC were syntactic rules the categories they introduced could be assigned structures by other rules in such a way as to place an undesirable element in the relevant position. In other words, there would be no way of preventing future instantiations of rules (whether basic or formed by metarule) from violating the conditions on cliticization.

For example, since clitics do not change the syntactic category of what they attach to, if a sentence like (3.87b) was formed on the basis of a sentence like (3.87a) the resulting form wanna would still be a verb.

3.87) a. I want to leave.
b. I wanna leave.

As such there would be no straightforward way of preventing this verb from being expanded by the conjunction rules responsible for sentences like (3.86). This would result in the generation of ungrammatical sentences such as (2.9b):

2.9) b. *I don't need or wanna hear about it.

Similarly, if a sentence like (3.88a) could be formed by the PS rules then a sentence like (3.88b) could also be formed by freely instantiating the rules responsible for (3.88a) with the feature [NUL]:

3.88) a. I am to leave and Pita's to leave also.
b. *I am to leave and Pitas's also.

We encounter the same type of difficulty with unbounded dependencies as well.

In order to maintain a syntactic analysis of AR and CC we would have to give up the view that PS rules are unordered as well as the idea that PS rules are node admissibility conditions, both of which are fundamental assumptions in GPSG. As a result we would be left with a less restrictive and much weaker theory. If we wish to preserve the theory as it is, we are forced to treat cliticizations as something distinct from syntax. The facts in a GPSG approach could not be otherwise without seriously altering its underlying claims. Thus we see that a theory of grammatical organization that has been argued for on independent grounds by many others falls out automatically if we employ a GPSG syntax.

Footnotes

*I would like to thank a number of people for their comments on various versions of this work. They are (in no particular order): David Dowty, Arnold Zwicky, Mike Geis, Brian Joseph, Rob Fox, Greg Stump and the members of Dowty's 1982 seminar on phrase stucture grammars. Any remaining errors are, of course, my mother's fault. I would also like to thank Karen Goldman and Isa Soto for kindly not finishing their theses before I finished mine. Erhard seems to think that he too deserves special mention.

1. Among others: Zwicky (1977), Klavans (1980), Zwicky (1982), Zwicky and Pullum (1982), Kaisse (1983a,b), and Pullum and Zwicky (forthcoming).

2. As developed in Gazdar (1981, 1982); Gazdar, Pullum and Sag (1982); Gazdar and Pullum (1981); Gazdar, Klein, Pullum and Sag (1982).

3. Klavans (1979) claims that another such distinction is that clitics always attach outside of any inflectional endings the host may have, arguing that apparent cases of endoclisis mentioned in Zwicky (1977) actually result from the clitic itself bearing suffixes.

4. Perlmutter (1970), George and Toman (1976), Klavans (1980).

5. Klavans argues that classical Greek provides examples of (1.7b) with so-called "stranded proclitics" such as ou in (i):

i. pôs gàr oú?
"for why not"

However, as Klavans herself points out (p.144) due to pre-pausal stress rules there is no way to tell if the "clitic" is truly attached to the following sentence. Since the element does have stress and does stand on its own, it is worth questioning in what sense it is a bound dependent in such sentences.

6. It should be noted that certain persons and tenses of these auxiliary verbs cliticize more freely than others. For the most part I will restrict myself to the forms is and has when discussing AR since they reduce most readily. In addition certain phonological considerations seem to discourage (though do not render impossible) AR. For a more complete discussion of the morphological and phonological factors involved in AR see Kaisse (1983a).

7. It should be pointed out that some dialects of English have more stringent restrictions on AR than others. These usually refer to the syntactic category of the phrase containing the host rather than the category of the word actually receiving the clitic. Therefore, even in conservative dialects, a variety of elements can serve as host.

8. An exhaustive list being: aspectual go, aspectual used, necessitative got, necessitative have, ought, suppose, and want in the sense of desire (rather than lack).

9. The fact that the [əz]/[ɨz] forms appear, at first glance, to have a wider distribution with reduced is/has than with the plural, third

person or possessive may simply be a result of there being phonological reduction rules with this output as well. Thus we have to distinguish between the two sources for these forms. In fact, in my own speech I prefer the vowelless alternate to the [əz]/[ɨz] forms at slow rates in sentences like:

i. John is going.
ii. Pete is going.

At faster rates [əz]/[ɨz] is acceptable, indicating that its ability to occur in positions not predicted by the general allomorphy rule under discussion is the result of a phononogical reduction rather than AR.

10. As argued for by Riemsdijk and Williams (1981).

11. While I find sentences like (2.16) perfectly grammatical there are sentences with Negative Constituent Presposing which do not sound as good:

i. ?Never's he to darken my door again.

I think this is because preposing with never is extremely stilted for me even without cliticization. Thus I would disagree with Kaisse's claim that there are no stylistic effects whatsoever involved in AR.

12. A possible response to this criticism would be that for purposes of AR the gap left by subject Wh-Movement "counts" as the first constituent in the sentence. Thus the reduced auxiliary in (2.22) and (2.3a,b) would still be a 2P clitic. However, such an analysis could not give a natural account of the presence of voicing assimilation in (2.22) or (2.3a) and would also reintroduce the notion of "gap" into Kaisse's analysis, something she had argued is not necessary.

13. This is not entirely clear from her article.

14. For a fuller discussion of the drawbacks to Selkirk's analysis see Pullum and Zwicky (forthcoming) and works cited therein.

15. This observation is due to Schmerling (unpublished manuscript). Note, however, that Postal and Pullum (1978) claim that a few speakers do accept sentences like (2.34b) (the one example they cite is Terry Langendoen), though they admit that they themselves find these sentences ill-formed. I am not aware of any other speakers who accept such sentences.

16. Sag and Klein also point out that rule numbers can be eliminated entirely in favor of indices on lexical categories. Furthermore, if you assume that metarules only apply to ID rules that introduce lexical categories and that numerical indices are contained in the feature matrices of lexical categories, then the claim that these indices are preserved under metarule application follows automatically. This is the approach adopted in some of the most recent GPSG articles.

17. See Gazdar and Pullum (1982) for a more complete description of the GPSG feature system.

18. In order to assure finiteness of the set of categories no feature is allowed to take itself as a coefficient (Gazdar and Pullum (1982)).

19. FOOT features are sometimes referred to as "binding features" as well.

20. Some versions of GPSG make use of a metagrammatical placeholder $H_n$, where $0 <= n <= 3$, to represent the head daughter in an ID rule. The HFC insures the identity of features between $H_n$ and its mother node. Gazdar and Pullum (1981) point out several advantages to this definition of head: 1) it allows the HFC to operate more generally in that it is responsible for all feature identity between mother and head daughter rather than just some of it, 2) the notion H simplifies the analysis of word compounds in English eliminating the need for parochial definitions of head in such cases and 3) H allows generalizations about English word order to be captured easily and without redundancies by the LP statements. For more details on the use of "H" see Gazdar and Pullum (1981, 1982).

21. This formulation of VP Deletion is slightly different from the one given in Gazdar, Pullum and Sag (1981) but more in keeping with the current approach to features. Since traces must retain their other category features there seems to be no need to have the trace dominated by VP[+NUL] as in earlier works.

22. This rule differs from the one given in GKPS in that it explicitly includes a trace element in its output.

23. In some current versions of GPSG this agreement is predicted by the Control Agreement Principle which says, simply, that verbs agree with their controllers. For more details see Klein and Sag (1982) and Gazdar and Pullum (1982).

24. These instantiations must also be consistent with any feature cooccurence restrictions (FCR's) which may apply, such as the FCR's which forbid a VP or A' from carrying a Wh feature.

25. Whether we want to give up this stipulation altogether is unclear. We may wish to have one class of metarules with this restriction and one without. If we do give it up entirely we will have to restrict the application of STM1 and STM2 appropriately, perhaps by reviving the Generalized Left Branch Condition.

26. Alternately one may wish to call the output of (3.45) something other than "S" to distinguish them from non-interrogatives; this is a minor detail.

27. This approach also entails a slightly different view of the organization of the grammar than the one taken by Sag and Klein outlined in section 3.1. Instead of ordering feature instantiation principles after metarule application, we must allow features to pass onto the ID doubles themselves in order to have S[+Q] expansions to serve as input to the rule in (3.45). Again, this is very similar to the treatment found in earlier versions of GPSG.

28. This version of the SAI metarule is based on the one given in Gazdar, Pullum and Sag (1981).

29. Given the organization of the grammar discussed in fn. 27 we would not even have to specify these features in the output of the SAI metarule. This is because 1) the HFC will insure that they are assigned to the V in the input rule and 2) features are preserved under metarule application unless otherwise specified.

30. The details of how such an FCR would work are not immediately relevant.

31. The discussion of so-called "Inversion" sentences presented here is greatly simplified and a number of details remain to be worked out. In particular, I avoid addressing the effects of definiteness and "heaviness" on linearization possibilities, cf:

i. a. Under this slab there is buried Joan of Arc.
   b. ?Under this slab there is buried Joan.
ii. a. Under this slab there is someone buried.
   b. *Under this slab there is Joan of Arc buried.

Such issues are beyond the scope of this work and in most cases do not critically affect the distribution of traces, which is our primary interest here. Also, it is not immediately clear whether we wish to maintain the distinction between existential there sentences, such as (3.66), and so called "presentational there" sentences, such as (3.67), suggested in Aissen (1975). These questions will be addressed in future work.

32. However one defines this class.

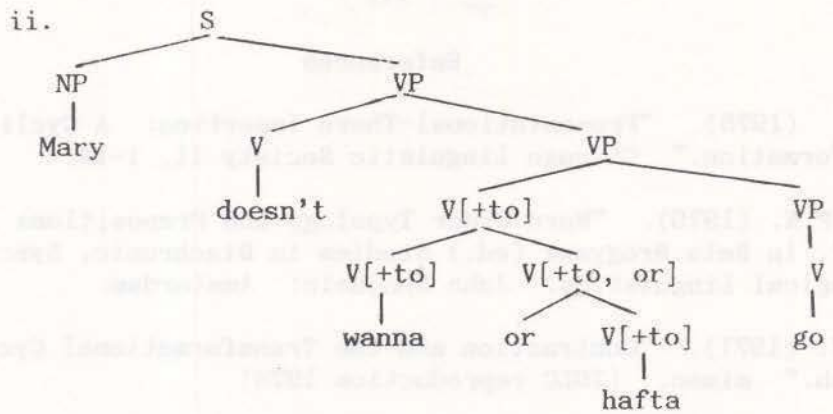
33. Again, I refrain from giving a formal rule for CC since it has been argued and will be argued again later that AR and CC are not and can not be syntactic rules in a GPSG system.

34. Exactly how these principles will be stated is somewhat problematic since there are certain restrictions on rightward dependencies that are not shared by leftward dependencies. Jacobson (1983 oral presentation, OSU) suggests that the non-unboundedness of rightward dependencies be captured by treating them as "double slashes" rather than single slashes, where double slash dependencies have the property of not being able to pass through certain types of nodes (bounding nodes). These issues, however, are outside the scope of this thesis; the matter that concerns us here is the placement of any trace elements, which would presumably be the same in both approaches.

35. Alternately, we could allow reduction to take place here with the result that the feature [+to] (or something to that effect) is assigned to the V. This feature will then trickle down onto each conjunct by general feature passing principles to produce sentences such as (i) with structures as in (ii):

i. Mary doesn't wanna or hafta go.

ii.




where a V[+to] is realized as its cliticized counterpart. Notice that the sentence in (i) will also have a possible source in:

iii. Mary doesn't want to or have to go.

Such an analysis can thus account for the feelings of some speakers that the sentence in (iii) requires "right node raising intonation" while the reduced sentence in (i) does not. Since (i) can also have the structure in (ii), it need not be analyzed as involving RNR. This is of course all quite speculative and depends upon adopting a particular view of what cliticization rules "look like". Whether this plan is feasible or not requires a great deal more study.

36. This point will be discussed further in the next section.

References

Aissen, J. (1975). "Presentational-There Insertion: A Cyclic Root Transformation." Chicago Linguistic Society 11, 1-14.

Anderson, P.K. (1979). "Word Order Typology and Prepositions in Old Indic", in Bela Brogyani (ed.) Studies in Diachronic, Synchronic, and Typological Linguistics. John Benjamin: Amsterdam.

Bresnan, J. (1971). "Contraction and the Transformational Cycle in English." mimeo. [IULC reproduction 1978]

Chomsky, N. (1976). "Conditions on Rules of Grammar." Linguistic Analysis 2, 303-351.

Chomsky, N. (1977). "On Wh-Movement", in P. Culicover, T. Wasow, and A. Akmajian (eds.) Formal Syntax. Academic Press: New York.

Chomsky, N. (1980). Rules and Representations. Columbia University Press: New York.

Chomsky, N. and H. Lasnik (1977). "Filters and Control." Linguistic Inquiry 8, 425-504.

Emonds, J.E. (1976). A Transformational Approach to English Syntax. Academic Press: New York.

Gazdar, G.J.M. (1981). "Unbounded Dependencies and Coordinate Structure." Linguistic Inquiry 12, 155-184.

Gazdar, G.J.M. (1982). "Phrase Structure Grammar", in P. Jacobson and G.K. Pullum (eds.) The Nature of Syntactic Representation. D. Reidel: Dordrecht.

Gazdar, G.J.M., E. H. Klein, G.K. Pullum and I.A. Sag (1982). "Coordinate Structure and Unbounded Dependencies". Stanford Working Papers in Grammatical Theory vol. 2. IULC: Bloomington

Gazdar, G.J.M. and G.K. Pullum (1981). "Subcategorization, Constituent Order, and the Notion "Head"", in T. Hoekstra, H. v.d.Hulst, and M. Moortgat (eds) Lexical Grammar II. Foris Publications: Dordrecht.

Gazdar, G.J.M. and G.K. Pullum (1982). Generalized Phrase Structure Grammar: A Theoretical Synopsis. IULC: Bloomington.

Gazdar, G.J.M., G.K. Pullum and I.A. Sag (1982). "Auxiliaries and Related Phenomena." Language 58, 591-638.

George, L. and J. Toman (1976). "Czech Clitics in Universal Grammar." Chicago Linguistic Society vol. 12, 235-249.

Jacobson, P. (1982). "Evidence for Gaps", in P. Jacobson and G.K. Pullum (eds) The Nature of Syntactic Representation. D. Reidel: Dordrecht.

Kaisse, E.M. (1983a). "The Syntax of Auxiliary Reduction in English." Language 59, 93-123.

Kaisse, E.M. (1983b). "The English Auxiliaries as Sentential Clitics", in Papers from the Parasession on the Interplay of Morphology, Phonology, and Syntax. Chicago Linguistic Society: Chicago.

King, H.V. (1970). "On Blocking the Rules for Contraction in English." Linguistic Inquiry, 134-136.

Klavans, J. (1979). "On Clitics as Words." Chicago Linguistic Society 15.

Klavans, J. (1980). Some Problems in a Theory of Clitics. Ph.D dissertation. University College London.

Klein, E.H. and I.A. Sag (1982). "Semantic Type and Control." Stanford Working Papers in Grammatical Theory vol. 2. IULC: Bloomington.

Lakoff, G. (1970). "Global Rules." Language 46, 627-639.

Lakoff, G. (1972). "The Arbitrary Basis of Transformational Grammar." Language 48, 76-87.

Nida, E.A. (1946). Morphology: The Descriptive Analysis of Words. University of Michigan Press: Ann Arbor.

Permutter, D. (1970). "Surface Structure Constraints in Syntax." Linguistic Inquiry 1, 187-256.

Postal, P.M. and G.K. Pullum (1978). "Traces and the Description of English Complementizer Contraction." Linguistic Inquiry 9, 1-29.

Pullum, G.K. and P.M. Postal (1979). "On an Inadequate Defense of Trace Theory." Linguistic Inquiry 10, 689-706.

Pullum, G.K. and P.M. Postal (1982). "The Contraction Debate." Linguistic Inquiry 13, 122-138.

Pullum, G.K. and A.M. Zwicky (forthcoming). The Syntax-Phonology Interface.

Riemsdijk, H. and E. Williams (1981). "NP-Structure." Linguistic Review 1:2, 171-217.

Sag, I.A. and E. Klein (1982). "The Syntax and Semantics of English Expletive Pronoun Constructions." Stanford Working Papers in Grammatical Theory vol. 2. IULC: Bloomington.

Selkirk, E.O. (1972). The Phrase Phonology of English and French. PhD dissertation. MIT.

Wood, W. (1979). "Auxiliary Reduction: A Unified Account." Chicago Linguistic Society 15, 366-377.

Zwicky, A.M. (1977). On Clitics. IULC: Bloomington.

Zwicky, A.M. (1982). "An Expanded View of Morphology in the Syntax Phonology Interface." Preprints for the XIIIth International Congress of Linguists, Tokyo.

Zwicky, A.M. and G.K. Pullum (1982). Cliticization versus Inflection: English n't. IULC: Bloomington.